Laryngeal reinnervation: feasibility studies and development of trial outcome measures

Marina Mat Baki

Thesis submitted for the degree of Doctor of Philosophy

University College London
The UCL Ear Institute

Declaration

I, Marina Mat Baki hereby declare that the work presented in this thesis is my own. Where information has been derived form other sources, I confirmed that this has been indicated in the thesis.

Abstract

The unifying theme of this thesis is a series of research studies that collectively amount to a feasibility study for clinical trials of laryngeal reinnervation for the treatment of vocal fold paralysis. The question 'Does laryngeal reinnervation or thyroplasty give better voice results for patients with unilateral vocal fold paralysis (UVFP)?' remains outstanding; a question that ideally requires a randomised control trial. However, randomised control trials in surgery face inherent surgeons' equipoise and recruitment issues that may lead to its failure. I performed a national survey of UK ENT consultants exploring their perception and obtaining crude numbers of eligible UVFP patients under their care for such trial, which revealed that the majority of ENT surgeons are receptive to the trial and the size of the potential patient pool is promising. I interviewed eligible UVFP patients to explore issues around the recruitment process, and this suggested that the proposed trial is feasible. Some phraseology used during recruitment that needed changing was identified, which may optimise the recruitment process for a trial.

In voice surgery trials, outcome measures should be multidimensional and standardised. Acoustic analysis has been proposed but has limitations. OperaVOX is a potential new acoustic analysis software developed to resolve some of these factors. I demonstrated that OperaVOX is statistically comparable to the 'gold standard', Multidimensional Voice Programme, for most principal phonatory outcome measures.

Another outcome measure- video-laryngostroboscopy, allows visual evaluation of characteristics and vibratory pattern of vocal folds. It is typically subjective that

requires inter- and intra-rater reliability study. Here, I demonstrated that certain parameters depicted substantial inter- and intra-rater reliability. However, I showed that rater training is required to improve the reliability of other parameters.

I investigated MRI as a potential non-invasive method to evaluate vocal muscles' denervation and reinnervation. I found that signal changes on the T2-weighted MRI larynx images correlated with electrophysiological results with good repeatability. Another MRI sequence, dynamic contrast enhanced- and diffusion weighted MRI, suggested reduced perfusion in paralysed muscles, whilst cine-MRI for vocal fold mobility assessment demonstrated considerable potential as a method to grade vocal fold mobility.

Finally, I present a small prospective case series of non-selective and selective laryngeal reinnervation in UVFP and unilateral vagal paralysis following vagal tumour excision respectively concomitant with injection laryngoplasty. Voice improvement was demonstrated by voice handicap index-10 and other multidimensional outcome measures, and these were supported by laryngeal electromyography and T2-weighted MRI outcomes. To my knowledge, this is the first multidimensional prospective study of laryngeal reinnervation and also the first to suggest that 3T MRI may be a promising outcome measure for future reinnervation trials.

In summary, I have shown that a randomised trial of laryngeal reinnervation versus thyroplasty is feasible in the UK, and have validated patient- and observer-rated outcome measures. I have also shown that MRI may offer an alternative to

electromyography in the assessment of laryngeal neuromuscular function in future trials and the clinic.

Acknowledgement

First and foremost, I would like to extend my sincere gratitude to my supervisors Martin Birchall and Shonit Punwani. They had tirelessly given their continuous support, guidance, advice and help throughout my research. I learned a lot from them either directly or indirectly how to be a good researcher. Their helpful comments for my thesis writing are greatly appreciated. Without them I couldn't have survived the PhD research course.

I would also like to express great appreciation to John Rubin, Anne Schilder and Guri Sandhu for their invaluable support and feedback throughout my research and patients recruitment. Thanks to Peter Clarke, Alasdair Mace, Elfy Chevretton, Khalid Ghufoor, Taran Tatla and other ENT surgeons for collaborating in my study.

I am thankful to the Centre for Medical Imaging of University College London team members: Shonit Punwani; David Atkinson; Alex Menys; Georgekutty Naduvilethil; Nicola Stevens; and Lorna Smith for assisting me in developing the MRI protocol and scanning the participants. The MRI study would not be possible without their help, support and extremely useful feedback.

I thank Anil Alexander, Oscar Forth and Owain Hughes for introducing me the OperaVOX, a new voice analysis software package. I was given an opportunity to test the prototype of the software of which subsequently I purchased and used the final version of the software package for my PhD research.

A heartfelt special thanks to Gary Wood, Philippa Ratcliffe and Marion Alston, for they had given unconditional commitment in evaluating video clips of the video-laryngostroboscopy study and collaborating in the OperaVOX study. I really appreciate time and technical support on acoustic analysis provided by Gary.

I am extremely thankful to Robert Royston, Bryan Youl and the Neurophysiology team members of Royal Free Hampstead Hospital. Robert had provided technical support for laryngeal electromyography tests and Bryan contributed his expertise in evaluating the electromyography tracing.

I thank JW Fairley and ENT UK Members for their voluntary contribution in the survey of perception of the proposed randomised control trial of laryngeal reinnervation versus thyroplasty.

I am very grateful to all staffs of Royal National Throat Nose Ear Hospital, University College Hospital and University College London. They are very kind, friendly and helpful throughout my research period. Anne Oliphant, Heather Fitzke, Angela Constantinou and Jumoi Abiona are some of the names that I could mention here.

My gratitude also extended to all volunteers of University College London and University College London Hospital PhD and masters students and staffs, and patients for participating in my study and willingly attended all the clinic visits and tests.

Thanks to the Otorhinolaryngology Department of Universiti Kebangsaan Malaysia Medical Centre especially to Abdullah Sani for infusing motivation and Muhammad Razif M. Yunus and my colleagues for their continuous moral support.

I would like to thank the Ministry of Higher Education Malaysia, National University of Malaysia (University Kebangsaan Malaysia) and Universiti Kebangsaan Malaysia Medical Centre for providing financial sponsorship, and Radiological Research Trust for funding the MRI research.

Last but not least, My deepest gratitude goes to my husband, Muhammad Nazri and children, Muhammad Zaim, Nurun Nuha and Sofiyyah Hani for their support, patience and understanding during my difficult times through out my study.

Table of content

List of figures

List of tables

Abbreviation

ADC	Apparent diffusion coefficient
AUC-1	Area under the curve at 1 minute
BVFP	Bilateral vocal fold paralysis
Cine-MRI	Cine magnetic resonance imaging
CT	Cricothyroid muscles
DCE-MRI	Dynamic contrast enhanced magnetic resonance imaging
DWI	Diffusion weighted imaging
ENT	Ear, Nose and Throat
F0	Fundamental frequency
LCA	Lateral cricoarytenoid muscles
IA	Interarytenoid muscles
LEMG	Laryngeal electromyography
MDVP	Multidimensional voice programme
ME	Maximum enhancement
MFR	Mean flow rate
MPT	Maximum phonation time
MRI	Magnetic resonance imaging
MUAP	Motor unit action potential
NHR	Noise-to-harmonic ratio
PCA	Posterior cricoarytenoid muscles
PIS	Participant information sheet
ROI	Region of interest
SoE	Slope of enhancement
SRI	Stroboscopic research instrument

TA	Thyroarytenoid muscles
T2-MRI	T2 weighted magnetic resonance imaging
UVFP	Unilateral vocal fold paralysis
VFaP	Vocal fold abduction potential
VFPa	Vocal fold phonation asymmetry
VFRa	Vocal fold respiration asymmetry
VHI	Voice handicap index
VOS	Voice outcome survey
VRQOL	Voice related quality of life

Chapter 1 Introduction

1.1 Intrinsic laryngeal muscles and the innervation

The larynx is important for speech, breathing and swallowing. It consists of a cartilaginous framework and muscular structures supporting a pair of vocal folds. The vocal folds adduct to the midline to make a firm contact during phonation and abduct away from the midline during breathing to allow movement of air during inspiration and expiration. Intrinsic laryngeal muscles that act as an adductor are the lateral cricoarytenoid (LCA), interarytenoid (IA), thyroarytenoid (TA) and cricothyroid (CT) muscle (Figure **1-1**). For abduction, the vocal folds are mainly controlled by posterior cricoarytenoid muscle (PCA).

The TA is a paired muscle forms the main mass of the true vocal folds. It arises from the lower half of the inner part of the angle of thyroid cartilage and inserts into the base and anterior surface of the arytenoid cartilage. It runs parallel with the vocal ligament and adherent to it. The TA is critical in phonation as vocal fold tension and mass are generated by its action. Changes in vocal fold tension affect the pitch of phonation while changes in vocal fold mass affect the glottic closure and indirectly changes the vocal intensity (Crumley, 1994a).

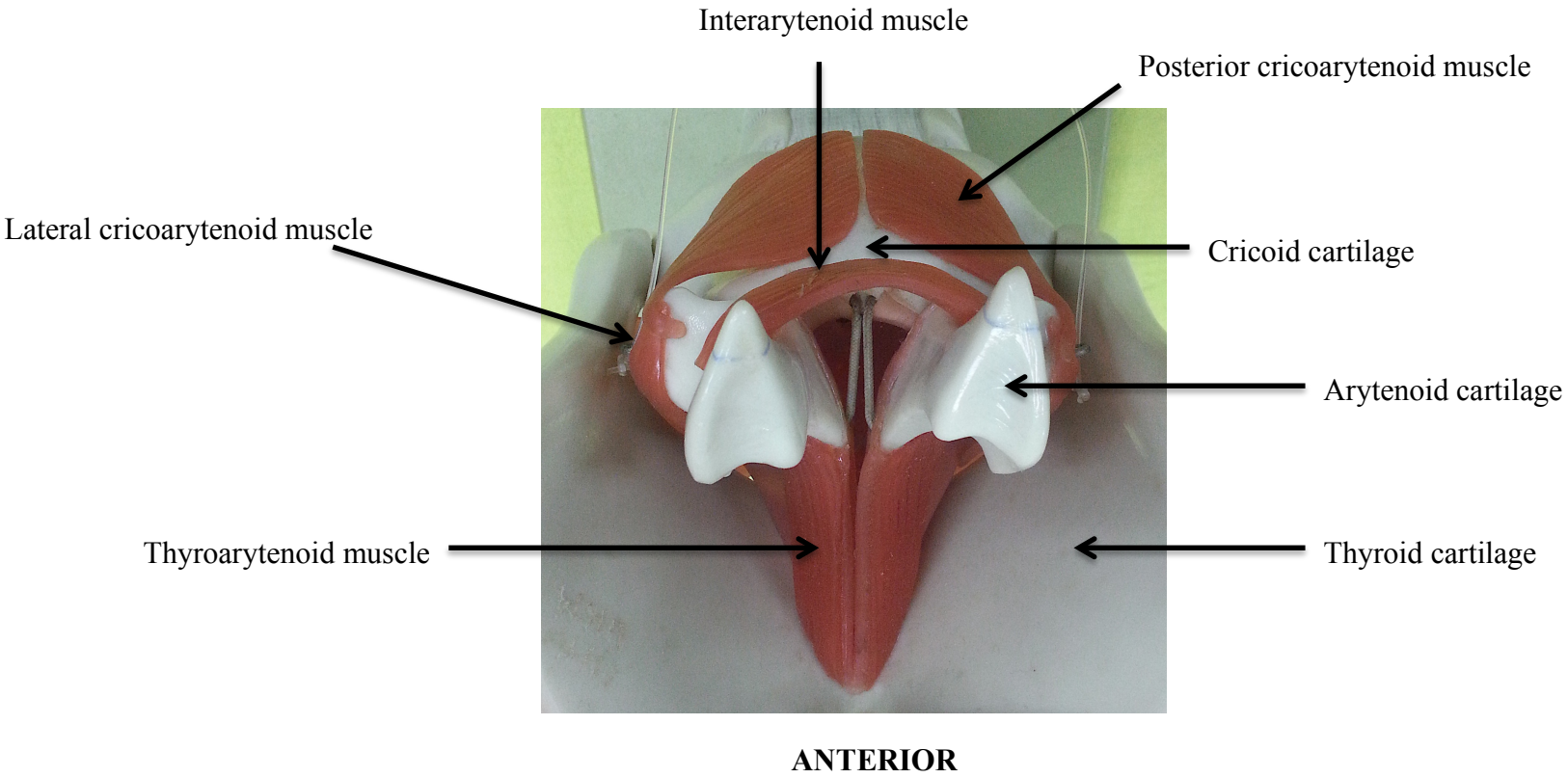


Figure 1-1: This picture shows a bird's eye view of intrinsic laryngeal muscles and the attachments.

The LCA is the most important vocal fold adductor and arises from the lateral border of the cricoid and inserts into the muscular process of the arytenoids. The contraction of LCA pulls the muscular process of the arytenoids in an anterior and medial direction. The vocal processes are pulled toward each other in an inward and downward movement thus bringing the vocal folds to the midline, closing the membranous glottis.

The IA is an unpaired muscle that has two types of fibres, transverse portion and oblique portion. The transverse portion runs horizontally across the posterior part of the arytenoid cartilages. The oblique portion runs obliquely from the base of one arytenoid to the apex of the other arytenoid. When the IA muscle contracts, it glides the arytenoids medially towards each other hence closing the posterior glottis. Some of the muscle fibers attach to the posterolateral surface of the arytenoids. Hence its contraction will also glide the arytenoids laterally thus stabilising the vocal folds during adduction (Mossallam et al., 1987).

The CT is a paired muscle that consists of two sets of muscle fibres, the pars recta and pars oblique. The pars recta originates from the lateral part of the cricoid cartilage and runs upward at an angle and inserts into the inferior border of the thyroid cartilage. The pars oblique also originates from the lateral border of the thyroid cartilage running obliquely upward and inserts into the anterior surface of the inferior horn of the thyroid cartilage. The CT muscle acts as a pitch changer by elongating and tensing the vocal fold. During contraction, the thyroid cartilage is tilted anteriorly and downward toward the cricoid cartilage hence increases the

distance between the anterior commissure of the thyroid and the arytenoid cartilages. The increase in distance stretches and elongates the vocal folds, decreasing their mass per unit of area and increasing the longitudinal tension.

The PCA is a paired muscle that abducts the vocal folds during breathing. It arises from the posterior part of the cricoid lamina and inserts into the muscular process of the arytenoid. Its contraction during high-pitched phonation stabilises the arytenoids, opposing the anterior force by the TA muscles and the vocal ligaments. Acute denervation of the PCA muscle causes partial subluxation of the arytenoid anteriorly, inferiorly and medially.

The intrinsic laryngeal muscles are innervated by two branches of the vagus nerve, the recurrent laryngeal nerve (RLN) and the superior laryngeal nerve (SLN). All the intrinsic laryngeal muscles- LCA, TA, IA and PCA, are innervated by the RLN except the CT muscle innervated by the SLN. Complete transection of the RLN completely denervates the ipsilateral muscles except for IA, which has crossed innervations. Complete transection of the vagus nerve affects both SLN and RLN hence symptoms following the injury may be worse in vagal nerve palsy compared to isolated recurrent laryngeal nerve palsy (Crumley and Izdebski, 1986; Woodson, 2007).

However, animal and human cadaveric studies on intrinsic laryngeal muscles innervation showed nerve supply of the larynx is complex. The classic description that all intrinsic laryngeal muscles are supplied by the RLN except the CT muscle which is exclusively innervated by the external branch of SLN (eSLN) is inaccurate (Maranillo et al., 2003; Martin-Oviedo et al., 2011). TA and PCA muscles were

shown to have dual innervation from the RLN and eSLN (Maranillo et al., 2003; Nasri et al., 1997) with the RLN predominating (Hydman and Mattsson, 2008). Hydman and Mattsson demonstrated electrophysiological evidence of the RLN and eSLN innervation of the PCA muscle in rats with an amplitude ratio of 4:1 (Hydman and Mattsson, 2008). Oviedo studied human laryngeal nerve connections in 13 patients undergoing total laryngectomy found that in 7 patients, the CT muscles were not exclusively innervated by the eSLN but received some innervation from the RLN as well (Martin-Oviedo et al., 2011). And in 3 patients, the PCA muscles were showed to receive some innervation from the eSLN. The dual innervation was contributed by either inter-connection between the RLN and eSLN or direct innervation by the distal branch of the eSLN (Martin-Oviedo et al., 2011; Paskhover et al., 2014).

The RLN comprises abductor and adductor motor neurons that are separated in the brainstem's nucleus ambiguus. Abductor motor neurons receive input from the inspiratory centre are located in ventral division whereas adductor motor neurons receive input from the cortex through the midbrain are located on the dorsal division (Flint et al., 1991; Gacek, 2001). This separation was demonstrated by injecting retrograde axoplasmic tracers such as horseradish peroxidase (HRP) in individual laryngeal muscles. However, the motor axons of abductor and adductor neurons are mixed when they enter the peripheral nerve pathway in the vagus as well as the RLN in humans.

1.2 Pathophysiology of peripheral nerve injury and repair

A nerve trunk consists of axons, axonal sheath, funiculus (bundles of fascicles) and epineurium (Figure 1-2). There are 3 layers of tissue arranged concentrically about the axon. Schwann cells form a single layer on the surface of the axon. It is then covered by 2 connective tissues, neurilemma (inner) and endoneurium (outer). Axons and the nerve sheath comprise a nerve fibre. A number of nerve fibres bundled in a sheath of connective tissue (perineurium) form a funiculus. A number of funiculi are held together by epineurium and form a nerve trunk.

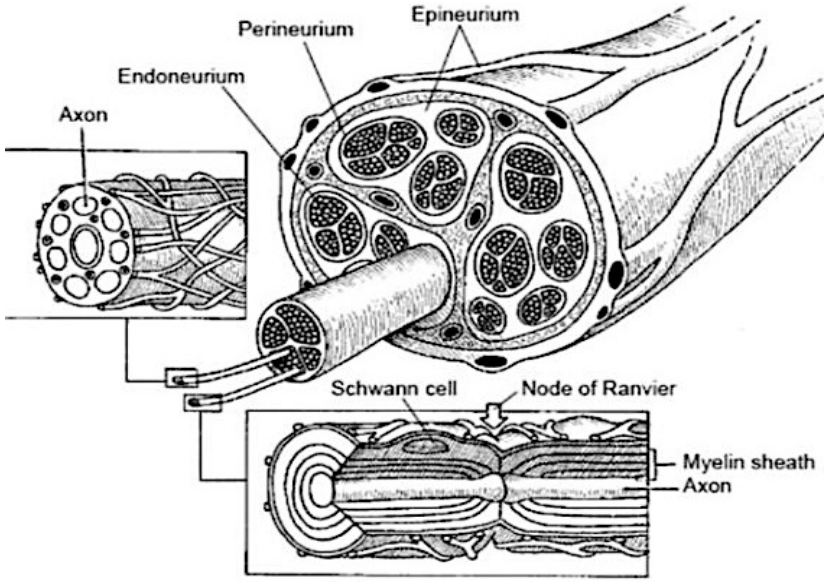


Figure 1-2: This diagram shows components of a nerve trunk (Biazar et al., 2010).

An essential component of nervous system is the nerve cell body with its associated dendrites and axons. Motor axons are closely associated with Schwann cells that are organised in a longitudinal sequence and meet at nodes of Ranvier where they are interdigitate with cellular processes. Electrical impulses pass from node to node as extracellular ions gain access to the axon.

Peripheral nerve injury was classified by Seddon (Seddon, 1942) and then was further expanded by Sunderland (Sunderland, 1951) as summarised in Table 1-1. The peripheral nerve injury was classified in 5 types according to its severity in which Wallerian degeneration occurs in type 3 to 5 nerve injury. Wallerian degeneration is a condition that occurs when the axon is separated from the cell body. The stump that was disassociated from the cell body will undergo degeneration followed by myelin degeneration and macrophage infiltration, which removes degenerated axons and myelin. Following Wallerian degeneration, the degenerated axons of peripheral nerves are able to regenerate due to the ability of the Schwann cells to promote a permissible environment. In type 3 to 5 injury, muscle end organ function may not return to normal despite effective regeneration due to the phenomenon of synkinesis.

Table 1-1: Sunderland classification of peripheral nerve injury

Type of injury		Wallerian degeneration
Type I (neuropraxia)	Conduction block	No
Type 2 (axonotmesis)	Axonal injury	No
Type 3	Type 2 + endoneurium injury	Yes
Type 4	Type 3 + perineurium injury	Yes
Type 5 (neurotmesis)	Type 4 + epineurium injury	Yes

1.3 Spontaneous laryngeal reinnervation and synkinesis

Injury to the vagus nerve or its branch of the recurrent laryngeal nerve (RLN) causes vocal fold paralysis. Common aetiologies of vocal fold paralysis include iatrogenic, neoplastic and idiopathic. Other less common causes are traumatic, aortic aneurysm, radiation induced and cardiovascular pathology. However, the commonest cause is iatrogenic including thyroid surgery, carotid endarterectomy, anterior approaches to the cervical spine, and heart or great vessel surgery (Rosenthal et al., 2007). Thyroidectomy has the highest incidence and the trend may still be increasing despite improvement in surgical skill and use of intra-operative RLN monitor (Benninger et al., 1998; Chen et al., 2007; Rosenthal et al., 2007; Yumoto et al., 2002). The injury could be temporary or permanent. Risk of permanent RLN palsy following thyroidectomy has been estimated at 1-2% (Jeannon et al., 2009) but this may be a considerable underestimate due to problems of follow up and reporting.

Following peripheral nerve injuries, denervated skeletal muscles undergo morphological changes. In chronic denervation, reports on the laryngeal skeletal muscle morphology in animal studies were inconsistent. Some documented atrophy and fibrosis in denervated laryngeal muscles (Shindo et al., 1992; Zealear et al., 1994) whereas others indicate that laryngeal muscles are resistant to atrophy for up to one year (Kano et al., 1991). Shindo et al. studied chronological electrophysiological, morphology and histochemical changes in a dog laryngeal nerve damage model and found that denervation atrophy and fibrosis occurred early at 2 to 3 months after nerve transection. After 3 months, a variable degree of partial reinnervation was observed. Large muscle fibres in clusters among small atrophic fibres were seen at 9 months indicating spontaneous reinnervation.

Studies suggest that denervated laryngeal muscles have a strong propensity of spontaneous reinnervation (Chen et al., 2011; Woodson, 2007). The reinnervation may be due to either (a) sprouting from surrounding intact nerve or (b) regeneration of the transected nerve. Woodson's study (2007) supported the theory of regeneration of the transected nerve whereby axons were found scattered in surrounding fibrous tissue and/or organised myelinated axons were detected in the regenerated nerve across the gap. Likewise, Lewis et al. found no direct contact between normal and denervated PCA muscles microscopically and in control animals and no identifiable nerve between the right and left experimental muscles (Lewis et al., 1991).

Spontaneous regeneration of an injured RLN in human is unpredictable and not universal. Chen et al. performed a cross sectional histologic study of injured RLN of unilateral vocal fold (UVFP) patients more than 6 months after injury who underwent laryngeal reinnervation surgery. Intraoperative exploration revealed that, of 29 cases, the RLN was transected in 24 cases whereas in five cases, the RLN stump exhibited continuity between the proximal and distal parts. Microscopic examination of all injured nerves revealed that the proximal and distal ends were connected by connective tissue. Histology demonstrated a varying degree of RLN axonal regeneration. The maximum number of axons found in the injured RLN was not more than 40% of axons in healthy control nerves taken from 15 laryngeal carcinoma patients undergoing total laryngectomy. The regenerated axons were randomly arranged in a manner that might suggest aberrant reinnervation (Chen et al., 2011).

Aberrant regeneration of abductor and adductor motor axons associated with a phenomenon of mass movement of vocal folds is called laryngeal synkinesis (Crumley, 2000; Dalgic et al., 2013; Flint et al., 1991; Woodson, 2007). This is hypothesised to be the main reason for absence of vocal fold mobility despite a robust spontaneous reinnervation of denervated vocal fold muscles. Woodson demonstrated that spontaneous reinnervation in TA was more robust than PCA and that this may indicate preferential reinnervation in the adductor than abductor muscles. Evolution of laryngeal synkinesis was depicted in rats undergoing reinnervation by end to end anastomosis immediately after a transection (Pitman et al., 2011). The rats were sacrificed at 4, 8, 12, 16 and 20 weeks and their histology was evaluated. Electromyography and vocal fold mobility were also evaluated. Vocal fold abduction was observed 4 weeks after the procedure and there was no vocal fold motion detected further in all but one animal. This phenomenon was shown to be due to synkinesis that was first detected at 8 weeks.

In humans, Blitzer et al. documented 50% incidence of laryngeal synkinesis in vocal fold paralysis. A careful diagnosis of laryngeal synkinesis based on laryngeal electromyography (LEMG) tracing findings was proposed by Maronian et al. Adductor synkinesis may be diagnosed when there is recruitment of motor units of the TA during sniffing either equal to or more than the recruitment of motor unit during phonation. The recruitment has to be ‘equal or more’ because it is common to detect TA activity on the LEMG during sniffing to stabilise the arytenoids and stiffen the vocal folds to avoid air turbulence (Hillel, 2001). Abductor synkinesis may be categorized as the presence of any significant recruitment of motor units in the PCA

muscle during phonation (Maronian et al., 2004). Therefore, a lower incidence of laryngeal synkinesis – 10%, than previously published was documented.

1.4 Vocal fold paralysis

1.4.1 Unilateral vocal fold paralysis

In the acute phase of UVFP, the affected vocal fold is flaccid and incapable of closing the glottis during speech. The vocal folds cannot make a firm contact to produce a strong voice. The voice is frequently breathy and easily fatigued (Crumley, 1994a; Misono and Merati, 2012; Woodson, 2007b). This problem may also lead to ineffective protective mechanisms by the larynx during swallowing causing aspiration that subsequently may lead to aspiration pneumonia. Following the acute phase that may last for several weeks, there will be slight improvement of voice as results of physiological compensation by the contralateral normal vocal fold (Crumley, 1994a). Subsequently, the voice may get spontaneous favourable or un-favourable outcome (Crumley, 1994a; Hartl et al., 2005). This depends on the position of the immobile vocal fold that is varied by few factors such as the degree of spontaneous reinnervation, synkinesis and may be the dual innervation from eSLN. For favourable outcome, the vocal fold is in median or paramedian position, without much vocal fold atrophy and nicely compensated by the contralateral normal vocal fold. This results in complete glottal closure during speech producing good voice. For un-favourable outcome, the vocal fold is in lateral position with evidence atrophy causing incomplete glottal closure during speech, producing breathy, easily tired and hoarse voice (Crumley, 1994a; Hartl et al., 2005).

Symptomatic patients with UVFP suffer reduced general health and quality of life due to social and emotional disturbance as a consequence of the voice disorder. In

one study, the mean of subscale scores of Assessment Medical Outcome Study Short-form 36 (SF-36) survey was significantly lower ($p<0.001$) from the normal population (Fang et al., 2008). The quality of life (QOL) measured with well-validated voice related quality of life scales - voice handicap index (VHI) (Jacobson et al., 1997), voice outcome survey (VOS) (Gliklich et al., 1999), voice related quality of life (VRQOL) (Hogikyan and Sethuraman, 1999) was improved significantly after surgical interventions (Hogikyan et al., 2000; Spector et al., 2001). Therefore, rehabilitation of the paralysed larynx is important to patients.

Rehabilitation of the paralysed larynx generally includes speech therapy during the first 6 months of the watch and wait period as 30 to 60% of patients with UVFP resolved spontaneously (Havas et al., 1999; Ramadan et al., 1998; Sulica, 2008) either due to spontaneous favourable voice outcome or recovery of vocal fold mobility. Where speech therapy alone is insufficient to improve the voice, surgical treatment is offered. The aim of surgical treatments for UVFP is to achieve optimum glottic closure during voice production. Various surgical treatment options have been proposed but, to date, there is insufficient evidence concerning the best surgical option to improve voice quality in patients with UVFP (Misono and Merati, 2012). The treatment options include injection laryngoplasty, Isshiki type I thyroplasty, arytenoid adduction and laryngeal reinnervation (Crumley and Izdebski, 1986; Isshiki et al., 1975, 1978; McCulloch and Hoffman, 1998).

Injection laryngoplasty is a simple procedure whereby a material is injected into the space lateral to the paralysed true vocal fold. This material increases bulk to the atrophied vocal fold and medialises it to the midline so it makes a firm contact with

the contralateral normal vocal fold to produce a strong voice. It is the least invasive procedure that avoids neck incision and can be performed either under general anaesthesia via transoral or local anaesthesia via transcutaneous. However it is normally recommended for patients with small glottal gap (Kwon and Buckmire, 2004; Lakhani et al., 1996) and the effect of the injection is commonly temporary for weeks or months as most of current materials resorbed with time. Therefore the patients may need repeated injections (Carroll and Rosen, 2011; Mallur and Rosen, 2010). Injection materials for injection laryngoplasty include fat, bovine gelatin, collagen-based products, hyaluronic acid, carboxymethylcellulose and calcium hydroxyapatite (Carroll and Rosen, 2011; Kwon and Buckmire, 2004; Mallur and Rosen, 2010; McCulloch et al., 2002; Wen et al., 2013). The clinical effectiveness of these injection materials varies from 4 to 6 months (Mallur and Rosen, 2010). There is no high evidence in the literature as yet to suggest which is the most ideal material to be used that provides long term effect (Lakhani et al., 1996).

Isshiki type I thyroplasty (Isshiki et al., 1975) is widely used to rehabilitate the permanently paralysed larynx (Leder and Sasaki, 1994; McLean-Muse et al., 2000). It is an operation in which an implant is inserted through a window in the thyroid cartilage lamina of the larynx. This implant medialises the paralysed vocal fold in the midline position allowing the normal opposite vocal fold to make firm contact and produce a stronger voice (Figure 1-3). However, theoretically voice improvement may not be long term due to progressive atrophy of the denervated laryngeal muscles. And there may be persistent mild hoarseness due vocal fold stiffness caused by the silastic block (Maronian et al., 2004; Uloza et al., 2005). Furthermore, this

procedure does not restore vocal fold tension, meaning that pitch variation remains suboptimal. Patients who had undergone thyroplasty also may require a second operation such as revision thyroplasty, injection laryngoplasty or arytenoid adduction to further improve the voice or to treat the deteriorated voice due to progression of denervation atrophy (Anderson et al., 2003).

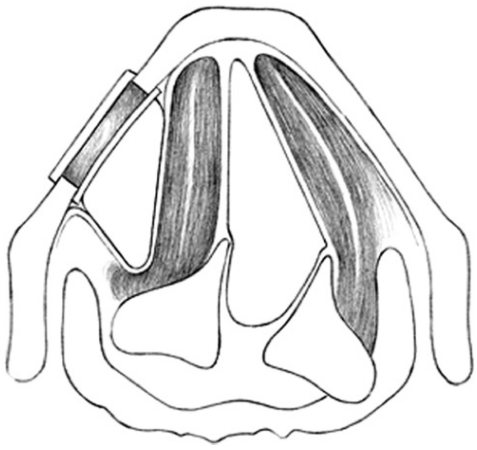


Figure 1-3: This diagram shows an implant inserted through the thyroid cartilage lamina window to medialise vocal fold in a thyroplasty operation (Franco Jr, 2012)

1.4.2 Bilateral vocal fold paralysis

For BVFP, both vocal folds are commonly unable to abduct during inspiration causing obstruction to the airway. It may present with critical impairment of respiration right through to simple voice changes (Benninger and Hseu, 2011; Cheung and McGinn, 2007; Dennis and Kashima, 1989). Treatments for BVFP are crucial and may be life saving. They are mainly surgery aiming to increase the glottal airway to alleviate the respiratory compromise. However, management of BVFP is a big challenge to laryngologists as one always needs to balance airway and voice.

Multiple surgical procedures had been developed in an attempt to improve sufficient airway patency. The first option of treatment in severe cases was usually a

tracheostomy. However, it is not frequently accepted by the patients as a long-term solution as there will always be a tube on the anterior neck that needs care and frequent cleaning. Surgical options to enlarge the glottal airway are laser posterior cordotomy, arytenoidectomy and laterofixation of the vocal fold with a suture and selective reinnervation (Cheung and McGinn, 2007; Dennis and Kashima, 1989; Gorphe et al., 2013; Marina et al., 2011a; Olthoff et al., 2005).

Laser posterior cordotomy and arytenoidectomy either partial or total involves opening the posterior part of the glottis. They have been preferred widely by ENT surgeons to enlarge the glottal airway (Cheung and McGinn, 2007; Gorphe et al., 2013; Olthoff et al., 2005; Segas et al., 2001). These surgical treatments have been reported as safe and effective in improving the airway but with the expense of hoarse, weak and breathy voice. Olthoff measured the airway resistance pre- and post-laser posterior cordotomy documented significant reduction in airway resistance and dyspnea ($p<0.001$). However the voice quality as measured by acoustic analysis was significantly worse ($p=0.001$). Voice related quality of life was significantly worsened by this type of operation (Yılmaz T et al., 2013).

Selective reinnervation is a procedure to improve the glottal airway by re-establishing the vocal fold abduction while preserving the vocal fold tone and bulk for good voice quality (Cheung and McGinn, 2007; Marina et al., 2011). Therefore it may be the choice to gain both airway improvement and preservation of normal or near normal voice quality. This topic will be discussed later in 1.7.

1.5 Laryngeal Reinnervation

A novel surgical approach for vocal fold paralysis is laryngeal reinnervation in which a foreign adjacent nerve is connected to a denervated muscle to re-establish working motor endplates and produce muscular contraction (Crumley, 1982). Reinnervation techniques described in the literature include nerve-to-nerve anastomosis (Crumley and Izdebski, 1986), nerve-muscle pedicle (Tucker, 1976) and nerve implant. Nerve 'anastomosis' is actually not an accurate term for nerve repair as it is supposed to be used in describing connection of two luminal structures such as bowel, blood vessels, bile ducts and others. The more accurate term is nerve repair. However nerve anastomosis is a common misuse in the literature related to reinnervation surgery.

The adjacent nerves that have been used for reinnervation are the ansa cervicalis, hypoglossal, superior laryngeal and phrenic. Nerve-to-nerve anastomosis involves connecting a donor nerve to the viable distal end of the RLN. The nerve-muscle pedicle technique uses a donor nerve attached to a small block of 'donor' muscle surrounding its terminal portion. The nerve implant technique is the modification of Tucker's technique whereby the nerve ending is implanted directly into the laryngeal muscle. Two main types of laryngeal reinnervation that have been reported are non-selective and selective.

Non-selective reinnervation is mainly for UVFP where providing tone and bulk to the laryngeal muscle is the primary concern whereas re-establishment of vocal mobility is secondary (Crumley, 1991; Lee et al., 2007; Lorenz et al., 2008; Wang et al., 2011). Selective reinnervation aims at re-establishing functional mobility to the vocal folds (Crumley, 1982; Fex, 1970; Marie et al., 1989; Rice, 1982). Successful

functional reinnervation to the PCA would allow BVFP patients to trigger vocal fold abduction during respiratory cycle (Marina et al., 2011; Remacle and Eckel, 2010), thereby overcoming the main handicap of this disorder.

1.6 Non-selective reinnervation - review

Non-selective laryngeal reinnervation presents some low morbidity operations with the potential to restore normal or near normal voice in UVFP patients without changing vocal fold pliability (Blumin and Merati, 2008; Crumley et al., 1988). The most common technique that had been used in the published literature is nerve-to-nerve anastomosis (ansa cervalis-to-RLN) as described by Crumley (Figure 1-5).

In this technique, the functioning ansa cervicalis nerve that overlies the internal jugular vein and the distal stump of injured RLN are identified and anastomosed without tension (Figure 1-4) (Crumley, 1991) . Wang et al. had proposed slight modification in which the main branch of the ansa cervicalis to be used instead of the branch to sternohyoid or omohyoid (Wang et al., 2011a). First, a horizontally oblique incision is made at the mid-thyroid cartilage level. The fascia between the strap muscles and the sternocleidomastoid muscles is opened to identify the ansa cervalis nerve superficial to the internal jugular vein. The main branch is identified and dissected to its termination in muscle. The RLN is then identified at the tracheo-oesophageal groove. In post-thyroidectomy cases, the RLN can be identified by doing intra-laryngeal dissection whereby the nerve will be found a few millimetres distal to the cricothyroid joint overlying the posterior cricoarytenoid muscle (PCA). The ansa cervicalis (proximally) and RLN (distally) are transected retaining

sufficient length for a tension free anastomosis, which is then performed using an operating microscope and 9/0 sutures, reinforced by fibrin glue**.**

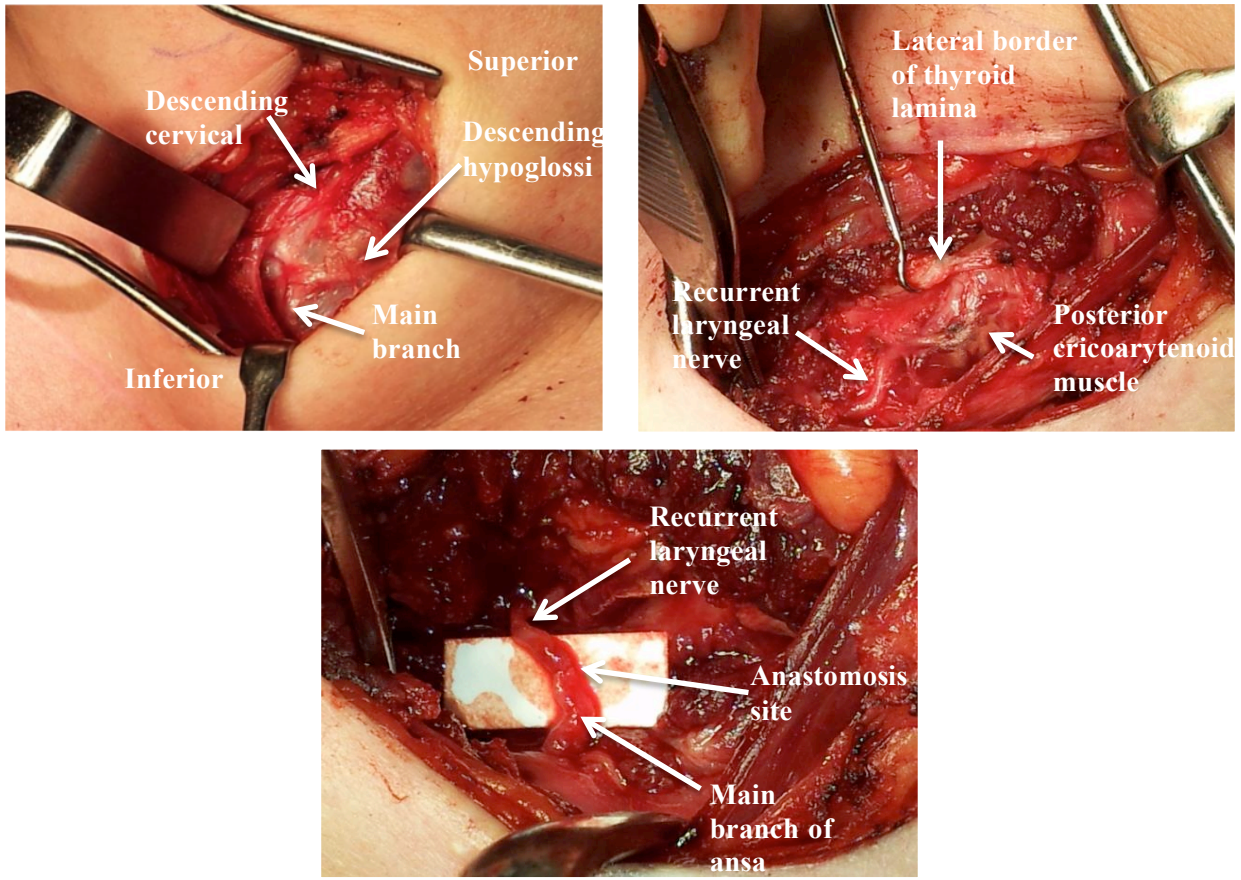


Figure 1-4: These intra-operative photos show: a) an identification of the ansa cervicalis nerve; b) an identification of the recurrent laryngeal nerve; and d) an ansa cervalis to recurrent laryngeal nerve anastomosis.

Quality of voice is determined by mass, elasticity, resistance and symmetry of the vocal folds, all of which are affected in UVFP. Denervation of the PCA and IA muscle causes anterior subluxation of the arytenoid and vocal process causing asymmetry of the vocal folds. Reinnervating the paralysed laryngeal muscles in UVFP employs the principle of gaining favourable synkinesis (Crumley, 2000) that will improve the resting tone and muscle bulk as well as restore the arytenoid to an upright position thus reducing the glottal incompetence (Crumley and Izdebski, 1986; Maronian et al., 2003). Firm contact of the two folds allows synchronous oscillation of the mucosa, reducing aperiodicity and perturbation of voice. There are several advantages of reinnervation over other surgical techniques.

This procedure avoids manipulation of the vocal fold therefore the mucosa remains supple and pliable. It precludes the need of implant adjustment to fine-tune the voice and this makes it suitable even in children. Furthermore, it preserves the possibility of other medialisation technique in future if it fails. Disadvantages of this technique are the requirement of a general anaesthetic, voice improvement is not immediate, requirement of an intact nerve donor and recipient and a suggestion of increased probability of delayed or failed reinnervation in older patients (Lee et al., 2007). Although the effect of reinnervation is delayed, the problems of aspiration and hoarseness may be assisted temporarily by bulking up the vocal fold with absorbable injectable substances.

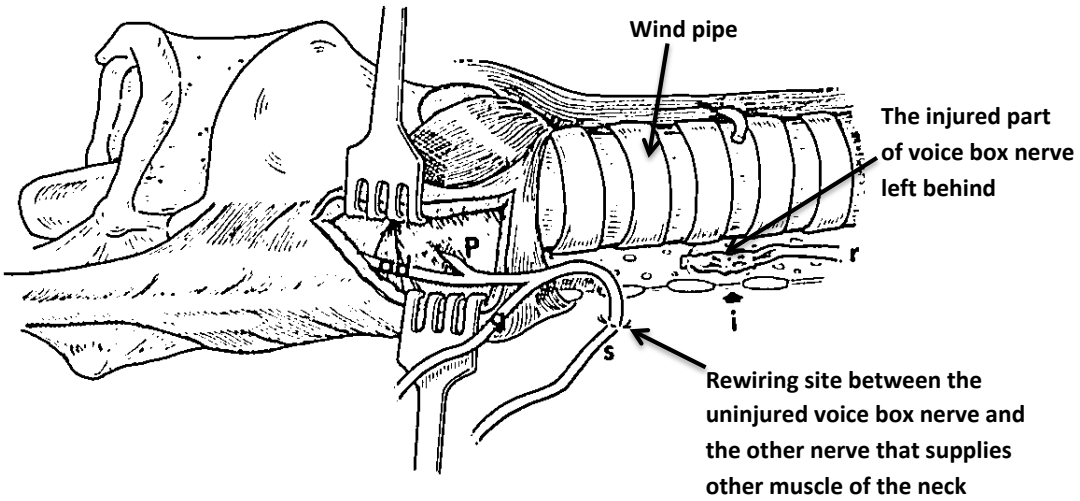


Figure 1-5: Diagrammatic picture of one form of laryngeal reinnervation- ansa-RLN operation (Crumley et al., 1988)

A recent systematic review of studies on reinnervation in UVFP patients was done to investigate the effectiveness in improving the voice (Aynehchi et al., 2010). The authors extracted 14 papers (Chou et al., 2003; Crumley, 1991; El-Kashlan et al., 2001; Lee et al., 2007; Lorenz et al., 2008; Maronian et al., 2003; May and Beery, 1986; Miyauchi et al., 2009; Olson et al., 1998; Paniello, 2000; Smith et al., 2008; Su et al., 2007; Tucker, 1989; Zheng et al., 1996) out of 686 papers retrieved from literature databases that met the inclusion and exclusion criteria. They were all of level IV evidence. Reinnervation techniques used in the studies were varied. Eight used ansa-RLN technique, most combined with injection laryngoplasty to assist the voice while waiting for the effect of reinnervation. Outcome measures used in the studies as well as the measurement time points were also heterogeneous. This review revealed that reinnervation is effective in improving voice to a varying degree. However due to the low quality of the current literature and heterogeneity in the outcome measures used that preclude comparison between studies, a formal prospective trial was recommended using standardised, internationally accepted outcomes. Here, I performed an additional review to update the 2010 findings.

1.6.1 Methods

Search engines were Pubmed, Cochrane and Scopus database from 2009 until 2013. Keywords used in carrying out this review were “vocal fold paralysis” OR “vocal cord paralysis” OR “recurrent laryngeal paralysis” AND “laryngeal reinnervation” OR “nerve anastomosis” OR “nerve-muscle pedicle” OR “nerve implant”.

Inclusion citeria:

Papers on non-selective laryngeal reinnervation in UVFP patients reporting the clinical outcomes that include case series, prospective or retrospective cohort, case control and randomised controlled studies.

Exclusion criteria:

1. Experimental studies of laryngeal reinnervation in animals
2. Other language than English
3. Selective laryngeal innervation in BVFP patients
4. Laryngeal reinnervation in children
5. Studies that did not include clinical outcome measures.

1.6.2 Results

From 44 papers retrieved, three papers fit the inclusion criteria were extracted. The papers reported a prospective study with and without control subjects, and an RCT (Paniello et al., 2011; Wang et al., 2011; Yumoto et al., 2010). Two studies (Marcum et al., 2010; Zur, 2012) were excluded due lack of clinical outcome measures. Another three studies (Hassan et al., 2011; Li et al., 2013; Wang et al., 2011) were excluded because the same population of patients was included from other published studies (Table 1-2). The three studies included for review are summarised in Table 1-3.

Yumoto et al. included 22 UVFP patients in a retrospective study that used a modified nerve implant technique combined with arytenoid adduction. The patients were followed up short term (1 to 6 months) and long term (7 to 36 months) post-operatively. Treatment outcome measures were maximum phonation time (MPT), maximum airflow rate and acoustic analysis (jitter, shimmer, harmonic to noise ratio and pitch range). The study revealed that there was a significant improvement in all outcome measures between pre-operative and post-operative review, at short term and long term time points. There was no data on neurological status of the vocal fold muscles included in the study. Two of 22 patients had difficulty in breathing necessitating tracheostomy to secure the airway. One patient who had surgical history of oesophagectomy, developed inspiratory dyspnoea possibly due to laryngeal closure reflex triggered by oesophageal regurgitation. Another patient had stridor due to laryngeal oedema. Both of them were decannulated in a few days post-surgery.

Paniello et al. included 24 patients in the data analysis of an RCT of reinnervation (combined with injection laryngoplasty) versus medialisation (thyroplasty with or without arytenoid adduction). The primary outcome measure was perceptual evaluation by untrained listeners (RUL). The number of patients was small due to premature termination of the study by the sponsor as a consequence of a low recruitment rate and some issues around obtaining informed consent and study administration. The study demonstrated a significant improvement between pre- and post-operative data of both groups (For RUL, medialisation group- $p=0.008$; reinnervation group- $p=0.027$). However, some data interpretation was not clear. They found that there was no significant difference of RUL, perceptual evaluation by

professionals (GRBAS score) and voice related quality of life (VRQOL) between the two groups. MPT improvement was better in the medialisation group that showed a significant difference at 6-months ($p=0.008$) and 12-months ($p=0.023$). A subgroup analysis, although methodologically inappropriate, purported to suggest that those aged less than 52 years had better results in reinnervation than thyroplasty: they achieved a perfect GRBAS score and normal range for VRQOL. LEMG results of the reinnervation group based on Koufman classification revealed that the neurological status was either improved or the same except in one patient. None of them had total denervation after the reinnervation procedure (Koufman and Walker, 1998). Three patients of each group needed additional surgery, injection laryngoplasty or thyroplasty, to improve the voice further or alleviate aspiration.

Wang et al. published a large retrospective study that included 237 UVFP patients undergoing ansa-RLN reinnervation and 237 normal control subjects. Post-operative outcomes were measured at 2 to 12 years. The outcome measures employed in the study were MPT, mean airflow rate, acoustic analysis, GRBAS score, videostroboscopy evaluation and motor unit recruitment on LEMG. The study demonstrated significant improvement in all parameters between pre- and post-reinnervation and no significant difference between the post-reinnervation voice quality and normal control subjects. There was no paradoxical vocal fold mobility observed. For GRBAS score, the inter-rater reliability was >0.76 and intra-rater reliability >0.81. The LEMG results revealed significant improvement in voluntary motor unit recruitment between pre- and post-operative measurement ($p<0.001$). The failure rate of the ansa-RLN reinnervation was 2% in which 4 of 237 patients did not achieve satisfactory vocal function. The causes of failure were postulated to be

inadequate removal of scarred end of the injured recurrent laryngeal nerve distal stump, interruption of recipient and donor nerve connection following evacuation of post-operative haematoma and long denervation course of more than 3 years that may contribute to insufficient laryngeal reinnervation. The complication rate was 5.5% (13 of 237 cases) and the most common were mild complications such as ecchymosis (11 cases). Other documented complications were haematoma, wound issues and stridor. Two patients had stridor due to haemorrhage that necessitate tracheostomy.

Table 1-2: All studies of human laryngeal reinnervation from 2009 to 2013

Author	Year	Study design	Level of evidence
Marcum et al.	2010	Case report	4
Yumoto et al.	2010	Retrospective	4
Hassan et al.	2011	Retrospective	4
Wang et al.	2011	Retrospective with normal controls	3
Paniello et al.	2011	RCT	2B
Wang et al.	2011	Retrospective	4
Zur et al.	2012	Retrospective	4
Li et al.	2013	Retrospective	4

Table 1-3: Summary of results of the three studies extracted in the systematic review

	Yumoto et al.	Wang et al.	Paniello et al.
MPT	**Pre**, 5.2(2.1); **post**, 13.1(5.1), *p<0.001* and continue to improve up to 36 months (*p<0.01*)	**Pre**, 6.18(4.87, 7.63); **Post**, 17.22(14.44, 21.78), *p<0.001*. No difference between the UVFP and normal subjects	MPT in ML significantly longer than reinnervation at 6 months (*p=0.009*) and 12 months (*p=0.023*)
MFR, ml/s	**Pre**, 713(680); **Post**, 145(58), *p<0.01* and remain unchanged up to 36 months *p=0.69*	Not done	Not done
Pitch range, semitone	**Pre**, 7.6(4.84); **Post**, 14.0(4.2), *p<0.01* and remain unchanged up to 36 months *p=0.69*	Not done	Not done
Jitter, %	**Pre**, 7.67(4.84); **Post**, 2.08(1.25), *p<0.001* and continue to improve up to 36 months *p<0.05*	**Pre,** 1.52(1.13, 2.04); **Post**, 0.33(0.21, 0.58), *p<0.00*. No difference between the UVFP and normal subjects (*p value ranged from 0.11 to 0.667*).	Not done
Shimmer, %	**Pre**, 12.28(7.17); **Post**; 6.7(4.14), *p<0.01* and remain unchanged up to 36 months *p=0.5*	**Pre**, 9.33 (7.95, 11.82); **Post**, 4.22 (2.59, 6.42) *p<0.001*. No difference between the UVFP and normal subjects	Not done
Noise measurement	HNR, **Pre**, 5.42(3.27); **Post**, 8.00(1.64) (*p<0.01*) and remain unchanged up to 36 months	NHR, **Pre**, 0.147 (0.113, 0.212); **Post**, 0.018 (0.012, 0.029) *p<0.001*. No difference between	Not done

	p=0.96	the UVFP and normal subjects	
RUL	Not done	Not done	No siginificant difference between the groups.
GRBAS	Not done	**Pre**, each item, ranged from 1.0 to 2.2; **Post**, 0, *p<0.001*	No significant difference between the groups
Patient self-evaluation	Not done	Not done	VRQOL, no significant difference between the groups
Videostroboscopy	Not done	Significant improvement of glottic closure, vocal fold position, vocal fold edge, and phase symmetry and regularity (*p<0.001*)	Not clearly shown
LEMG	Not done	Significant improvement in voluntary motor unit recruitment between pre- and post-operative measurement (*p<0.001*)	Koufman classification was either improved or the same from before the operation except in one patient. None of them had total denervation.

1.6.3 Discussion

Significant long-term voice improvement was documented in UVFP patients when the results were analysed between pre- and post-operative. Ansa-RLN technique as described by Crumley was still the most common technique used. Wang et al. had improved the technique by using the main branch of the ansa cervicalis that has more axons to regenerate to the injured RLN instead of ansa hypoglossi to sternohyoid, and they described details of the surgery for a successful reinnervation. Successful results were supported by evidence of reinnervation on LEMG data in the two studies that used ansa-RLN technique. Yumoto et al. employed modified Tucker's technique and combined it with an arytenoid adduction. Their study had shown significant immediate and long-term voice improvement but lack of evidence of the neurological status to support the voice improvement (Yumoto et al., 2010).

Of the 3 studies, Wang et al. had included the largest number of patients. In this study, the success rate of ansa-RLN technique was found to be high with only 2% of failure to achieve satisfactory vocal function. The rate of major complications involving airway compromise was low (2 of 237 patients). None of these patients developed paradoxical vocal fold mobility (Wang et al., 2011).

Treatment outcome measures still vary between the studies and do not follow the suggested basic protocol (Dejonckere et al., 2001). The only parameter that was measured in all three studies was MPT. Nevertheless, GRBAS score and acoustic analysis (jitter, shimmer, NHR) were measured in two studies. Self-patient evaluation that reflects patients' satisfaction of voice improvement and its psychosocial impact is important but was only measured in one study. Paniello et al. chose evaluation by untrained listeners as the primary outcome measure, which is to

my knowledge the first study to use it. It was also the only study that stated the primary outcome.

There was only one RCT since the last review in 2009. The study showed that there was no significant difference in the voice quality evaluated by untrained and trained listeners' as well as self-perceptual evaluation between reinnervation and medialisation group. However the study was underpowered since the sponsor terminated it prematurely due to inherent recruitment issues. Therefore, there remains the central question as to which treatment is better for UVFP – reinnervation or thyroplasty. Thus, a further RCT is required but must be properly planned and piloted. Issues around patients' recruitment need to be explored before embarking the trial to help ensure feasibility.

1.7 Selective reinnervation - review

The amended part of this review has been published by the Current Opinion of Otolaryngology & Head and Neck Surgery (Marina et al., 2011).

Research and literature on selective laryngeal reinnervation are limited down to the variability and complexity of nerve supply. Airway problems resulting from simultaneous bilateral RLN transection in animal studies are another factor impeding progress. Thus, most studies on humans relate to non-selective reinnervation in UVFP.

There are numerous animal studies in the literature exploring the feasibility of selective reinnervation of the PCA to reacquire vocal fold abduction. Many of the studies showed good results, confirmed by endoscopic observation of vocal fold abduction either spontaneous or by electrical stimulation, the electromyographic characteristics and histological evidence.

Results of studies on humans are sparse and based on case series. In the earlier reports, nerve muscle pedicle techniques using ansa hypoglossi-omohyoid muscle pedicle were applied to reinnervate the PCA muscles (Applebaum et al., 1979; Tucker, 1976). Vocal fold abduction was observed at four to eight weeks post operatively. However, there was no report on the same technique for bilateral reinnervation in human since then. The other technique used was the split-phrenic nerve–graft procedure that however failed to return any abduction to the vocal fold even though the results in animal studies were encouraging (Crumley, 1983).

Marie has developed a bilateral selective reinnervation technique for bilateral vocal fold paralysis patients after a series of animal studies and had promising results. He reinnervated the PCA with right upper phrenic nerve root with the help of an

interposition free nerve graft and the adductors with thyrohyoid branch of the hypoglossal nerve (Figure 1). He started a clinical prospective trial in humans in 2003 and to date more than 40 patients have received this technique (unpublished data). He reported a preliminary result of his early 12 patients (Remacle and Eckel, 2010). Four of six evaluable patients had decreased dyspnoea. Three of six patients achieved active arytenoids adduction and all are decannulated, albeit one of them after an arytenoidectomy.

Selective reinnervation to the adductor muscle in BVFP patients is important not for the return of adduction but to prevent atrophy and flaccidity of the vocal fold which may lead to voice deterioration. The flaccid vocal fold may also be sucked to the midline during inspiration and this would obstruct the airway. Furthermore it avoids the development of aberrant reinnervation of the adductor muscles from the nerve axons of the abductor muscle.

1.7.1 Phrenic nerve

Selective laryngeal reinnervation in BVFP patients is mainly to re-establish the vocal fold abduction during inspiration through reinnervation of the PCA muscle. The PCA muscle is the only laryngeal muscle that abducts the vocal fold. Therefore the adjacent nerve to be used for reinnervation must contain a homogenous composition of motor neurons that are active during the inspiratory phase of breathing. The most suitable nerve in this regards is the phrenic nerve (Baldissera et al., 1989). The PCA muscle is therefore selectively reinnervated with the phrenic nerve to allow the separation of re-growing homogenous axons in the phrenic nerve to the PCA muscles. The criticism of using the phrenic nerve is the possibility of diaphragmatic

paralysis after the nerve resection that may worsen the respiratory handicap especially if the laryngeal reinnnervation failed to improve the laryngeal obstruction. However, unilateral phrenic nerve resection has been shown to cause remarkably little effect on ventilatory capacity in patients with normal lungs (Fackler et al., 1967). Experiments on animals suggested that highest root of the phrenic nerve or the accessory phrenic nerve which is supplied by C5 (in man) could be used for reinnervation as it functionally preserves valuable diaphragmatic innervation and unilateral phrenic nerve resection has limited effect on the diaphragm (Marie et al., 1989, 1997)

Distribution of the phrenic nerve roots in animals has been described. The distribution in humans was first investigated to delineate the topographical distribution of C3 to C5 innervating the diaphragm by stimulating the nerve roots in three patients undergoing selective laryngeal reinnervation. This was the first reported study in man. The diaphragmatic activity was detected by surface electromyographic recordings. It was revealed that in humans, the C4 instead of C5 has a prominent proportion of phrenic nerve resembles the similarly large distribution of C5 in cats and rabbits. However, the study also indicated that C5 can provide a non-negligible proportion of the phrenic nerve that may have the possibility of a significant functional role of the accessory phrenic nerve in humans (Verin et al., 2011).

1.7.2 Variability of nerve supply to the PCA

The PCA muscle is critical to BVFP patients as it is the only abductor for respiration. The plausibility of reinnervating this muscle has been questioned since anatomical

studies confirm the variability of its nerve supply. The supply patterns to the human PCA muscle had been elegantly described by Maranillo et al. It was found that the main trunk of RLN divided into two terminal branches, the anterior and posterior division, just before entering the larynx with variable location from the cricothyroid joint (Maranillo et al., 2003). A PCA branch arose from the anterior branch, which subsequently gave off a number of branches forming a neural plexus under the deep surface of PCA muscle. Eller RL et al. confirm Maranillo's findings and added that the PCA muscle also receives nerves from the interarytenoid branch in 70% of larynges (Eller et al., 2009). Identification of PCA branches is not straightforward as a posterior window in the thyroid cartilage was required in 57.1% in male specimens and 69.2% in female specimens in a cadaveric study (Kwak et al., 2010).

The complexity of neural plexus, variability of nerve patterns and the involvement of nerve supply from the IA branch explain the occurrence of laryngeal synkinesis even following selective reinnervation of the PCA muscle. However this variability should not preclude selective reinnervation. The feasibility and efficiency had been proven in animal models (Kingham et al., 2005). A well designed surgical technique using this procedure has been described on human cadavers and also on patients with some encouraging preliminary results (Kwak et al., 2010; Remacle and Eckel, 2010).

1.7.3 Selective reinnervation in vagal nerve tumour patients

Cervical vagal nerve tumours could be classified into nerve sheath tumours and neuroendocrine tumours (Gilmer-Hill and Kline, 2000; Urquhart et al., 1994). The nerve sheath tumour is commonly known as vagal schwannoma and the neuroendocrine tumour as vagal paraganglioma. They are commonly benign in nature although there is 10% incidence of malignant tumour in vagal

paragangliomas. Most of patients with these tumours are asymptomatic as the progress is gradual and slow in nature. The commonest presentation is a neck mass. However lower cranial nerve palsies are possible as a result of tumour compression in which the commonest is vagal nerve paralysis (Urquhart et al., 1994).

Definitive treatment for vagal paragangliomas is excision surgery. It is usually associated with post-operative unilateral vagal paralysis manifesting dysphonia, dysphagia and aspiration. Additional operative complications of glossopharyngeal and hypoglossal nerve paralysis may worsen the swallowing problems.

Multiple paraganglioma is common especially in patients with positive family history (Urquhart et al., 1994). Netterville et al. reported 37% of vagal paragangliomas were bilateral (Fang et al., 2011; Netterville JL et al., 1998). Multicentricity of vagal paragangliomas is a significant factor to consider in selecting the therapeutic options (Lozano et al., 2008). Three of 19 patients with paraganglioma as reported by Urquhart et al. were subjected to irradiation therapy instead of surgery due to the potential of synchronous or metachronous multicentricity. This was designed to avoid the incapacitating morbidity of bilateral vocal fold paralysis (BVFP) with tracheostomy, an even greater burden for these patients whose swallowing may already be compromised (Netterville et al., 1993). Irradiation therapy has not been shown to cure paragangliomas and it has the drawback of possible temporal bone osteoradionecrosis and increased risk of malignant transformation (Netterville JL et al., 1998). Irradiation therapy also carries the risk of fibrosis to the vagus nerve that consequently results in vocal fold paralysis and makes later surgery more difficult.

Surgical rehabilitation to improve voice following unilateral vagal paralysis due to vagal paraganglioma excision that have been reported are thyroplasty, injection laryngoplasty and laryngeal reinnervation (Lamarre et al., 2010; Miller et al., 2000; Urquhart et al., 1994). Urquhart et al. documented voice improvement following Teflon injection laryngoplasty in all 11 of 16 patients with vocal fold paralysis following vagal paraganglioma excision surgery. Bielamowicz et al. reported early arytenoid adduction efficacy in 17 vagal paralysis patients after skull base surgery in which 76% of the patients had complete glottal closure with excellent voice quality as rated by speech therapists (Bielamowicz et al., 2000). Fang et al. mentioned great improvement in vocal fold performance following medialisation surgery but it was limited by the problem of nasal flow leakage that results in hypernasality (Fang et al., 2011). However, assessement of the real results of surgery in these studies are limited by inadequate published data.

For voice rehabilitation results following laryngeal reinnervation, Lee et al. included patients with vagal paraganglioma who underwent laryngeal reinnervation in their retrospective study. The study documented favourable outcome of ansa-RLN anastomosis in treating patients with UVFP characterized by improvement in patients' voice perception, acoustic analysis and laryngeal appearance (Lee et al., 2007). Lorenz et al. reported long lasting improvement of the glottic closure and maintenance of the vocal fold edge following laryngeal reinnervation at the same sitting as the actual vagal nerve tumour resection (Lorenz et al., 2008). However in these studies, voice outcome data analysis of paraganglioma patients was precluded

due to unavailability of the preoperative data, although the paper states that voice was normal before the tumour resection. Lamarre et al. reported a case of vagal paraganglioma resection with primary reinnervation of the larynx in the same sitting (Lamarre et al., 2010). In this case report, apart from ansa-RLN anastomosis, cricothyroid-to-cricothyroid reinnervation and greater auricular nerve-to-superior laryngeal nerve anastomosis were performed for completeness. The patient showed good glottic closure at one year postoperatively and had minimal problems with swallowing, requiring enteral tube feeding for 6 days.

Results of surgical rehabilitation by medialisation to reduce aspiration following unilateral vagal paralysis have shown variable outcomes. Thyroplasty and cricopharyngeal myotomy do not cure the problem (Miller et al., 2000). Arytenoid adduction is not helpful either as 66% of patients with unilateral vagal paralysis following skull base surgery had disappointing immediate effects of swallowing rehabilitation despite the excellent voice outcome. They required enteral feeding for more than six months. On the contrary, Flint et al. and Netterville et al. documented beneficial outcomes of medialisation thyroplasty in managing aspiration in this group of patients. Flint et al. reviewed the outcomes of thyroplasty in reducing aspiration in patients with dysphagia and aspiration following recurrent laryngeal nerve (RLN) paralysis or high vagal injury. 70% of patients with severe dysphagia had successfully discontinued enteral feedings (Flint et al., 1997). However duration of time taken for the discontinuation of the enteral feeding was not reported. The rest remained on enteral feeding with one of the patients undergoing a secondary operation of cricopharyngeal myotomy. Netterville et al. showed that thyroplasty expedited the consumption of initial oral intake and adequate oral intake with the

reduction in duration of hospital stay (Netterville et al., 1993). However, measures of swallowing and aspiration reduction in these studies were poorly recorded.

Significant issues of dysphonia, dysphagia and aspiration caused by unilateral vagal paralysis following vagal paraganglioma excision definitely require rehabilitation. The surgical rehabilitation is ideally performed at the same sitting to minimise post-operative morbidity. Injection laryngoplasty using current available substances mostly provides a temporary effect that requires repeated injections. Thyroplasty is commonly done under local anaesthesia since patients' cooperation is deemed necessary by most practitioners to fine-tune the voice according to implant size and placement. Laryngeal reinnervation may be a good option since the operation is preferably done under general anaesthesia and it does not require fine-tuning of voice. The surgical reinnervation that aim to re-establish tone and bulk of the denervated muscle may be better than other surgical techniques as Woodson in 2007 showed that spontaneous regeneration did not occur in vagus transection in cats possibly due to the relatively long course of the nerve to RLN to which only a small proportion of axons go. Furthermore, re-establishment of vocal fold mobility may be possible by performing selective reinnervation and this may save patients from irradiation therapy or tracheostomy should the tumor grow later on the opposite neck. A prospective study with standardised outcome measures at different time points is necessary to demonstrate the effect of selective reinnervation in unilateral vagal paralysis following the first operated side of vagal paraganglioma.

1.7.4 Conclusion

A potential surgical option to improve the airway and at the same time preserving the voice in BVFP patients as well as for UVFP following vagal nerve tumour resection may be selective laryngeal reinnervation. This surgical option does not preclude the irreversible laser surgery should it fail to improve the airway obstruction. It involves reinnervation of the abductor and adductor laryngeal muscles. A suitable surgical technique with an appropriate donor nerve needs to be employed to ensure success. The surgical technique must be able to address the issue of reinnervation to the PCA muscle to trigger abduction during respiratory cycle, the preservation of good voice by strengthening the adductor muscles as well as prevention of laryngeal synkinesis. However the evidence of success on humans in the literature is limited to date.

1.8 Randomised controlled trial in UVFP intervention– the need and problems

RCT is widely regarded as the best method to evaluate the efficacy of surgical interventions. However, not all types of surgical trials are feasible. There are three types of RCT in surgery; Type I - trials involving comparison between medical treatments in surgical patients, Type 2 – trials comparing surgical techniques and Type 3 – trials comparing medical treatment and surgical treatment (McCulloch, 2002).

Results of adequately powered RCTs are classified as level 1 (highest grade) in evidence-based medicine (EBM). Clinicians should integrate EBM in their clinical practice to ensure treatments delivered had been investigated regarding the safety and efficacy. A cross sectional survey of publications in otorhinolaryngology (ORL) journals in 1999 revealed that most therapeutic recommendations were based on level 4 evidence (Bentsianov et al., 2002). A retrospective review to investigate how many level 1 and 2 evidence were in general ORL between 2004 and 2009 showed that the high level of evidence was lacking compared to general medical literature publications. Level 1 and 2 evidence in voice literature is even rarer (4.6%) than in general ORL (31%) (Benninger et al., 1994). To date only three RCTs dedicated to the study of unilateral vocal fold paralysis (UVFP) treatments have been published (Hertegård et al., 2004; Lau et al., 2010; Paniello et al., 2011). There is a need for ORL surgeons and researchers to conduct more RCTs in this field.

Various postulations were given to justify the slow rate of RCTs in voice literature although no clear evidence was given. It may be secondary to small number of patients in certain pathological groups, lack of reliable objective measures and sparse

number of drugs treating voice disorders. RCTs in surgery face methodological and practical challenges greater than those of drug trials due to factors including varying surgeon's and patients' equipoise, learning curve of new surgical techniques, difficulties to secure funding, blinding issues and variations between centres in multicentre trials (Ergina et al., 2009; McCulloch, 2002; Young et al., 2008).

The RCT has a rigorous design to minimise biases. Components that are essential in a RCT to reduce the biases include blinding, standardisation of the study interventions and standardisation in patient selection. A standardisation of patient selection according to the outlined inclusion and exclusion criteria is necessary to minimise confounding factors. However a restriction of inclusion criteria may preclude the generalisability of the outcome of the study (Cook et al., 2003; McLeod, 1999). Unlike drug trials, standardising surgical techniques is difficult as surgeons may vary in their experience and preference of technique particularly if it involves complex surgical procedure and multicentre participation (Ergina et al., 2009; McCulloch, 2002; McLeod, 1999). Blinding is particularly difficult in surgical trials, hence Solomon and McLeod found that only a third of RCTs in surgery had adequate blinding (Solomon and McLeod, 1995). Patients and surgeons are difficult to be made unaware on which type of surgery has been allocated due to issues like different type and site of incision, between open and closed methods, and between general and local anaesthesia. Inability to blind the patients in surgical trials causes bias especially if the primary outcome is subjective improvement in disease spesific symptoms or quality of life.

Costs to run a RCT in surgery are expensive (Cook et al., 2003; McCulloch, 2002; Solomon and McLeod, 1995). Funding bodies may be reluctant to award the grant to surgical trials due to the issues and difficulties mentioned above and poor quality of previous research.

Another essential component of a RCT is, an adequate recruitment of patients according to the sample size calculation to meet the desired power and significance. Poor recruitment rate is a major issue in RCTs in surgery (McCulloch, 2002; Thoma et al., 2010). Factors that may contribute to this problem are lack of surgeons' and patients' equipoise. Equipoise is a condition whereby there is no preference between treatments in which both treatments are thought to be equal in outcome (Lilford et al., 2004).

Lack of surgeons' equipoise may be attributed to the strong preference and familiarity of certain surgical technique (McCulloch, 2002). This may lead to unwillingness to participate and recruit patients into the trial. Surgeons who agree to participate but had hidden preference towards a particular surgical technique may subconsciously contaminate the consenting process with phrases or terminologies that may affect patients' willingness for randomisation (Donovan, 2002; Donovan et al., 2009; Lilford et al., 2004; Mills et al., 2003). Surgeons who are positive about randomisation also may have conflict when they face a particular patient with special clinical characteristics (Donovan et al., 2014; Hamilton et al., 2013; Lilford et al., 2004). This in turn may contribute to the lack of patients' equipoise thus affect the recruitment.

Lack of patients' willingness to participate and be allocated to any surgical option arms is commonly due to strong patient treatment preference or lack of equipoise (Solomon and McLeod, 1995). The strong treatment preference may be attributed to the patient's own research on the network or information that is received from other sources such as primary doctors, therapists or nurses. Other factors are like logistical issues, additional demand of the trial, worry of uncertainty and concern about information and consent (Hamilton et al., 2013; Ross et al., 1999). Unlike drugs, surgery is largely non-reversible and invasive. Patients tend to be reluctant to participate as they want to have the power to decide which surgery they will receive.

The RCT is accepted as the most rigourous study design in experimenting effectivenes of certain surgical technique. However RCTs in surgery face many issues that may affect the feasibility of of the trial. Nevertheless the issues may be minimised by looking for strategies to overcome the methodologic difficulties and proper planning. Surgeons' and patients' equipoise issues lead to poor recruitment. It was recommended that perception of surgeons is explored and patients' recruitment is properly planned and piloted and hence a feasibility study that includes qualitative methods is necessary (Lovato et al., 1997).

1.9 Voice outcome measures

Voice quality assessments are optimally multidimensional due to complex nature of voice (Baylor et al., 2005; Carding et al., 2009a; Dejonckere et al., 2001). There is considerable diversity in treatment outcome measures with regards to functional voice assessment in UVFP used in the literature (Baylor et al., 2005). This hinders meta-analyses of the results of voice treatment. A basic protocol for pre- and post operative treatment outcome measures has been proposed to achieve uniformity in the voice research methodology and to allow relevant comparisons with the literature when presenting or publishing the results (Dejonckere et al., 2001). The components of basic protocol are acoustic analysis, voice-specific patient self-reporting outcome, voice perceptual evaluation by professional raters, video-laryngostroboscopy assessment and aerodynamic analysis.

Acoustic analysis is a non-invasive computer based objective assessment of voice quality by measuring the properties of recorded voice signals. A patient's voice saying vowel /a/ is recorded in a quiet or sound treated room and then played on a software programme to measure the acoustic parameters. The acoustic analysis is potentially useful in measuring treatment outcomes following any interventions for vocal fold pathologies. It measures irregularity of mucosal vibration of the vocal folds caused by change in mass, elasticity and symmetry that may affect the F0, jitter, shimmer and NHR. Therefore, it provides an objective measure of the efficacy of voice treatments (Carding et al., 2004) and may provide valuable feedback for patients. The common measurement includes fundamental frequency (F0), jitter, shimmer and noise-to-harmonic ratio (NHR) (Amir et al., 2009). Jitter, shimmer and NHR, are the most common parameters reported in voice research (Brockmann-

Bauser and Drinnan, 2011). F0 is a measurement of the mean frequency of mucosal vibrations of the vocal folds. Jitter and shimmer are amplitude perturbation measurements that measure cycle-to cycle frequency and amplitude variation respectively in the analysed voice sample. NHR is a measurement of degree of hoarseness obtained by estimating the proportion of noise in the subject's voice.

There are a number of validated voice-specific patient self-reporting tools in the literature. They are used to evaluate the impact of voice disorders to patients' life and daily activity. Voice Handicap Index (VHI), Vocal Performance Questionnaire (VPQ) and Voice Symptom scale (VoiSS) have been used widely in voice research. All of these tools were shown to have excellent internal consistency (Cronbach's coefficient = 0.81 to 0.95) (Carding et al., 2009). Of the three tools, VHI reached the highest repeatability with intraclass correlation value of 0.83. However VHI is a long questionnaire that consists of 30 questions. Rosen et al. had done further factor analysis, shortened the VHI to 10 questions (VHI-10) and validated it (Rosen et al., 2004). It comprises three domains: functional (5 items), physical (3 items) and emotional (2 items). Deary et al. compared VHI-10 and VPQ as both are short and convenient. It was found that both are internally consistent and good overall indicator of the severity of voice disorders (Deary et al., 2004). The VHI-10 has been used in other retrospective or prospective trials involving UVFP patients and was able to demonstrate treatment effect before and after the operations (Carding et al., 2004; Hogikyan and Rosen, 2002; Misono and Merati, 2012; Rosen et al., 2000).

Voice perceptual evaluation is a subjective assessment of voice quality during speech task by professional raters using a validated scale. In voice research, patients'

running speech voices are normally recorded and saved in audio files. The audio files are then anonymised and listened by a panel of raters who then grade the voice quality. The widely used scale for this purpose is GRBAS scale as proposed by Hirano. The G is for grade of the overall voice quality, integrating all deviant components. R is for roughness or harshness, B is for breathiness, A is for asthenia and S is for strain. The evaluation encompasses four point grading scale (0, normal or absence of deviance; 1, slight deviance; 2, moderate deviance; 3, severe deviance) (Hirano, 1981). Highest inter- and intra-rater reliability have been demonstrated for G, R and B (Carding et al., 2000; Webb et al., 2003).

Video-laryngostroboscopy is regarded as a vital assessment tool in voice clinics. It is widely used as a clinical tool to examine mucosal waves of vocal folds, glottal gap and vocal fold mobility. It incorporates a scope with an attached camera to visualise the vocal folds, which is connected to a stroboscope machine that enables evaluation of dynamic effect of vocal fold vibration- mucosal waves. In Greek, *Strobos* means whirling, while *scopein* means watching or observing (Faure and Muller, 1992). Normal vocal folds vibrate during phonation that ranged 70 to 500 Hz in males and 130-1,000 Hz in female (Švec and Schutte, 1996). Light from the stroboscopy that illuminate the vocal folds during vibration match the frequency of the voice with slight delay in few seconds. Thus travelling mucosal waves can be seen showing the vibratory pattern of the vocal folds. It gives an indication of the health of tissue situated between the superficial mucosa (cover) and the vocal muscle (body) (Faure and Muller, 1992; Hirano, 1974). Evaluation of video-laryngostroboscopy findings is typically qualitative: rating by voice experts assessing the visual findings of vocal folds- vocal fold edge, amplitude and symmetry of mucosal wave and glottal closure,

as first described by Hirano and Bless (Hirano and Bless, 1993). Research tools have been developed to improve the reliability of video-laryngostroboscopy evaluation (Poburka, 1999; Poburka and Bless, 1998; Rosen, 2005).

The simplest aerodynamic measurement of voicing is maximum phonation time (MPT). It is a measurement of maximum vocal performance that is assessed manually using a stopwatch to time the duration. The patient is normally asked to sustain vowel /a/ at a comfortable loudness of voice after taking a normal breathe. This task is repeated three times and the longest duration is recorded. A more advance measurement of aerodynamic of voice is by using Aerophone (model 6600, Kaypentax, Lincoln Park USA). It is a tool to measure air leakage during phonation that indirectly measures closure of glottic gap. Parameters that normally measured are- maximum SPL dB and maximum flow rate (MFR) at voicing L/s that are indirectly calculated from intra-oral articulatory pressures. Each patient is fitted with a standard anaesthesia face-mask covering the nose and mouth. A small tube that connects the machine and the patient's mouth is placed just slightly beyond the upper and lower incisors. The patient is then asked to say /pa//pa//pa/ for 3 times at minimum, comfortable and loudest loudness level. There should be no phonation break at each loudness level once data acquisition is initiated. The device is connected to a computer system with software that generate the data required.

There has been no consensus to date on the primary outcome measures for trials evaluating the efficacy of treatments for UVFP (Misono and Merati, 2012). To my knowledge, there were three RCTs in surgery for UVFP. Each of the three RCTs used different primary outcome measures which was self-reported visual analogue

scores of voice quality (Hertegård et al., 2004), the well-validated voice handicap index (VHI) (Jacobson et al., 1997; Lau et al., 2010) and perceptual evaluation of voice by untrained listeners (Paniello et al., 2011). A meta-analysis comparing voice outcomes between injection laryngoplasty and thyroplasty used VHI as the primary outcome measures. However, crucially, surgeons' choice of outcomes might not truly represent patients' concerns (Ergina et al., 2009). Therefore it is important to explore what is the main concern to UVFP patients about their voice to determine the primary outcome measure that is most likely to be meaningful to patients in a full trial.

1.10 Acoustic analysis

There is an increasing interest in using voice-specific patient self-reporting outcome to monitor treatment effect. It may be the most valuable as it reflects patients satisfaction and may be more meaningful (Carding et al., 2009; Hogikyan and Rosen, 2002). Although the common questionnaires used in voice literature have been validated, and were found to be reliable and sensitive to change (Carding et al., 2009), they are limited by subjectivity. Furthermore, in surgical trials, patients blinding is frequently impossible. Therefore, acoustic analysis is a useful outcome measure to support patients' self-evaluation improvement of voice.

The common measurement of acoustic analysis includes fundamental frequency (F0), jitter, shimmer and noise-to-harmonic ratio (NHR) (Amir et al., 2009). These parameters are measured from a recorded voice either on sustained vowel or continuous speech. The classic way of measuring these acoustic parameters is by analysing the sustained vowel /a/, /i/ or /u/. This method has been widely used because it is quick and simple. It is also preferred because the measurement of sustained vowels is relatively not affected by individual special speech characteristics like speaking rate, speaker's dialect, intonation and idiosyncratic articulatory behaviour (Krom, 1995; Maryn et al., 2009; Parsa and Jamieson, 2001). The acoustic values of this method also have been used as the standard values in previous studies. On the other hand, acoustic analysis of continuous speech may represent more 'real world' day-to-day communication context than the sustained vowels (Klingholtz, 1990; Moon et al., 2012) as voice disorder symptoms frequently emerge in conversational voice. However the analysis of continuous speech is more challenging due to the dynamic properties of voice signals and necessity of

demarcating voiced and unvoiced segments (Maryn et al., 2009; Parsa and Jamieson, 2001). Acoustic perturbation measures (jitter, shimmer and NHR) of voice signals of sustained vowels were found to be more accurate than of continuous speech (Parsa and Jamieson, 2001; Zhang and Jiang, 2008).

Previous studies and reviews suggested that the results of acoustic analysis- inter- and intra-software, except for F0, were moderately correlated (Carding et al., 2009a) and authors suggested that such measures should not be used independently to classify vocal fold lesions (Amir et al., 2009; Brockmann-Bauser and Drinnan, 2011). Nevertheless it may be still useful in measuring dysphonic voices in UVFP (Baylor et al., 2005; Inagi et al., 1997a), giving objective evidence of treatment effect before and after intervention. Acoustic analysis has been used widely in papers reporting significant voice improvement in UVFP following surgical intervention (Baylor et al., 2005; Chowdhury et al., 2013; Hassan et al., 2011; Ryu et al., 2012; Schneider-Stickler et al., 2013; Wang et al., 2011a).

The lack of reliability of acoustic analysis may be due to several confounding effects such as different algorithms used by different software programmes (Bielamowicz et al., 1996), recording environment, type of microphone, frequency and amplitude of phonation during recording and distance of the microphone from the lips (Brockmann-Bauser and Drinnan, 2011). Therefore in an attempt to employ acoustic analysis as a secondary outcome in a multicentre RCT, those confounding effects have to be minimised. Measurement of acoustic parameters that is done at a particular day and time may not represent the patient's overall voice especially if the voice varies through out the day. Performing acoustic analysis at different time in a day and in different days is difficult due to logistical reasons. A low cost software

application installed in a mobile Apple gadget that recently available may meet the specifications above (http://www.operavox.co.uk/wp/). However it has not been rigorously tested for its reliability and validity before.

1.11 Assessment of denervation and reinnervation of laryngeal muscles

1.11.1 Laryngeal electromyography

Laryngeal electromyography (LEMG) is a tool to assess neurological status of intrinsic laryngeal muscles. It requires an experienced laryngologist, electromyographer and a neurophysiologist to perform the procedure. LEMG is done while a patient lying on a couch with neck slightly extended. A laryngologist inserts a monopolar needle into the TA muscles through the cricothyroid membrane while an electromyographer operates the electromyography machine. A neurophysiologist then evaluates the LEMG dynamic tracings that normally done subjectively.

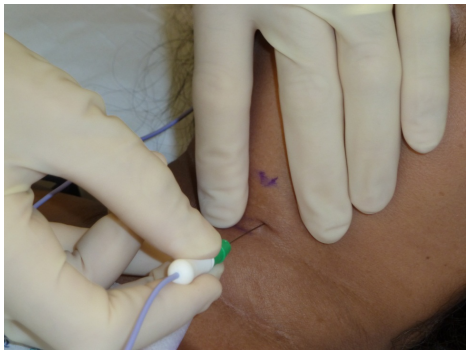


Figure 1-6: Laryngeal electromyography procedure. This photo shows a needle being inserted at the anterior part of the neck through the cricothyroid space to reach targeted laryngeal muscles.

LEMG provides information of normal or abnormal tracing of motor unit morphology, motor unit recruitment and spontaneous activity. Recruitment is defined as activation of motor unit during volitional activity like speech and sniffing during respiration. Abnormal tracing patterns are spontaneous activity (fibrillation potentials, positive sharp waves), polyphasic potential, nascent unit or decreased in motor unit recruitment. Vocal fold paralysis is confirmed when the abnormal patterns are seen that indicates denervation or reinnervation of laryngeal muscles. The

presence of polyphasic potential and/or nascent unit indicates reinnervation pattern whereas the presence of spontaneous activity indicates a denervation pattern. Following RLN injury, denervation and reinnervation patterns may be seen concurrently as spontaneous reinnervation is common (Bielamowicz and Stager, 2006; Chen et al., 2011; Koufman et al., 2001; Woodson, 2007a). Laryngeal synkinesis should be suspected if the LEMG results could not adequately explain vocal fold immobility.

LEMG has been used in clinical practice for diagnosis confirmation and prognosis evaluation of vocal fold paralysis, site and level of nerve injury, differentiation between paralysis and cricoarytenoid joint fixation and botox injection for spasmodic dysphonia (Anderson and Sataloff, 2004; Bielamowicz and Stager, 2006; Koufman et al., 2001; Mostafa et al., 2004; Munin et al., 2003; Rickert et al., 2012). Koufman et al. found from their retrospective LEMG data of 415 vocal fold paralysis patients that it altered the timing and type of surgical intervention in 40% of cases (Koufman et al., 2001). Ability of LEMG to provide prognosis of motion recovery varies due to qualitative nature of the assessment and presence of synkinesis. Rickert et al. had done meta-analysis of LEMG found that the positive predictive value was 90.9% whereas the negative predictive value was 55.6% indicating that it is excellent in predicting non-recovery of vocal fold motion but not in motion recovery. Absence or significantly decreased motor unit recruitment and presence of spontaneous activity usually indicate poor prognosis for motion recovery (Mostafa et al., 2004; Munin et al., 2003; Rickert et al., 2012). Fibrillation potential may be absent even in axonal loss probably due to the small number and size of the muscle fibers per motor unit of laryngeal muscles (Munin et al., 2003). Smith et al. did a study to improve the use of LEMG in predicting prognosis of vocal fold motion recovery by incorporating

quantitative analysis to the qualitative analysis of the LEMG in which the turns of amplitude per second were calculated. They found that the NPV increased to 89% and the PPV was 100% with sensitivity of 66.6% and specificity of 100% (Smith et al., 2012). Nevertheless, LEMG in clinical practice remains a qualitative tool.

LEMG is important in 'ruling in' vocal fold paralysis before reinnervation procedures. In non-selective reinnervation, apart from voice clinical test results, eligibility for the procedure is based on confirmation of paralysis and poor prognosis for vocal fold mobility recovery. Absence of motor unit recruitment and presence of spontaneous activity usually indicate the poor prognosis. Maronian et al. suggested criteria for reinnervation procedure when there is dense denervation detection, defined as 1 or 2 motor unit action potentials (MUAP) which are not activated during speech attempts or the presence of fibrillation potentials (Maronian et al., 2003). Polyphasic potentials indicate on-going reinnervation however, according to Koufman classification (James A. Koufman et al. 2001), polyphasic potential with absence of spontaneous activity indicates old injury which has stopped reinnervating.

LEMG is the only way to understand the physiologic response following reinnervation as histology evidence is lacking in human studies. In monitoring treatment effect of non-selective reinnervation, the return of laryngeal muscle tone is the main concern but not vocal fold mobility. The laryngeal muscle tone may be evaluated from the number of normal motor unit action potential (MUAP) and the recruitment of motor unit during voluntary laryngeal muscles contraction (Bielamowicz and Stager, 2006). Favourable outcomes of reinnervation may be indicated by the improvement of motor unit recruitment and presence of polyphasic

potentials as they were significantly more common in patients who had clinical recovery.

However, LEMG requires considerable operator skill, is invasive and muscles are not directly visualized during needle positioning. LEMG results have to be evaluated with caution if the patient is having significant pain due to the procedure as the recruitment may be artificially attenuated (Smith et al., 2012). It requires an experienced laryngologist, electromyographer and a neurophysiologist to get a reliable result (Volk et al., 2012). The otolaryngologist has to be someone who familiar with laryngeal anatomy and has been doing botulinum injection to the larynx by external approach regularly. For these reasons, I hypothesised that high resolution MRI may be a non-invasive alternative or adjunct to provide evidence of denervation and reinnervation in vocal fold paralysis patients.

1.11.2 Magnetic resonance imaging

MRI has been useful in demonstrating denervation changes of skeletal muscles affecting the morphology and physiology (Kamath et al., 2007; Koltzenburg and Yousry, 2007) and predicting prognosis (Yamabe et al., 2008). It may objectively depict evidence of denervation and reinnervation by measuring the muscle atrophy, signal changes on T2-weighted MRI (T2-MRI) (Bendszus et al., 2002; Kikuchi et al., 2003; Viddeleer et al., 2012; Yamabe et al., 2008), and perfusion changes on dynamic contrast enhanced MRI (DCE-MRI) and diffusion-weighted imaging (DWI).

In peripheral nerve neuropathies, a hyperintense signal on T2-MRI may be due to an increase in extracellular fluid secondary to oedema or increased in vascularity or fatty degeneration (Bendszus et al., 2004; Kikuchi et al., 2003; Polak et al., 1988) depending on the duration of nerve injury. The abnormal signal intensity changes on T2-MRI have been shown correlated with prognosis of nerve injury in rats (Yamabe et al., 2008) and with hand function in humans (Viddeleer et al., 2012). Diffusion-weighted imaging (DWI) is a type of imaging that study water motion in tissues by calculating the apparent diffusion coefficient (ADC) values. The principles of DWI are based on random movement of water molecules in tissues. It potentially can be used to investigate tissues perfusion and extracellular fluid space in denervated muscles (Viddeleer et al., 2012). DCE-MRI is a dynamic post-gadolinium contrast T1-weighted imaging method used to investigate tissue microcirculation as well as vascular changes in tissue following treatment as well as to diagnose benign or malignant tumours (Knopp, Michael V et al.). It is based on fast MRI sequences in cinematic display of certain tissue or lesion during and after intravenous gadolinium injection in which the distribution of the gadolinium is repeatedly imaged. This will give functional information related to vascular density, structure and permeability of microvasculature of a lesion. Benign angiogenic processes lead to increase in enhancement whereas the most distinct characteristic of malignant lesions is leakiness of the vasculature.

MRI of the larynx has been used in the investigation of tumours and aetiologies of vocal fold paralysis. However, its use in assessment of denervation and reinnervation of laryngeal muscles is absent from the literature to date. Sakai et al. investigated normal laryngeal structures in 62 participants using 1.5T machine and had shown

that MRI is efficient in demonstrating laryngeal muscles (Sakai et al., 1990). In vocal fold paralysis patients, MRI has been shown able to clearly depict denervation changes on the TA and PCA (Borges, 2010; Connor et al., 2006). A retrospective study was done to evaluate the ability of MRI in depicting denervation atrophy of laryngeal muscles (Romo and Curtin, 1999). This study involved 20 UVFP patients in whom four had had both CT scan and MRI. Denervation atrophy was detected with high frequency: 70 – 90%. They concluded that PCA muscle atrophy may be a useful indicator of RLN paralysis when other findings are subtle or equivocal. However there were no studies to correlate MRI changes with electrophysiological changes in denervated laryngeal muscles.

1.12 Assessment of visual findings of vocal folds

Laryngeal imaging is important for diagnosis of laryngeal disorders and has a role in assessing treatment outcomes. To date, various techniques have been used to evaluate phonatory function of vocal folds qualitatively or quantitatively including videostroboscopy (Omori et al., 1996; Rosen, 2005), videokymography and digital kymography to record vocal fold vibration (Verikas et al., 2009). Videostroboscopy in particular is widely available and regarded as a vital assessment tool in voice clinics. It is widely used as a clinical tool to examine mucosal waves of vocal folds, glottal gap and vocal fold mobility. Normal vocal folds vibrate during phonation that ranged 70 to 500 Hz in males and 130-1,000 Hz in female (Švec and Schutte, 1996). Light from the stroboscopy that illuminate the vocal folds during vibration match the frequency of the voice. Thus travelling mucosal waves can be seen showing the vibratory pattern of the vocal folds.

1.12.1 Visual findings of vocal folds' vibration patterns and characters

Evaluation of videostroboscopy findings is typically qualitative: rating by voice experts assessing the visual findings of vocal folds- vocal fold edge, amplitude and symmetry of mucosal wave and glottal closure, as first described by Hirano and Bless (Hirano and Bless, 1993). In UVFP, due to discrepancy in the vocal folds' vertical height, tensing capability and mass, and inability to make a firm contact with the opposite normal vocal fold, the mucosal waves are commonly asymmetrical with a significant glottal gap. Videostroboscopy findings have been shown to correlate with acoustic and perceptual analysis of voice (Uloza et al., 2005). It has been used to monitor treatment effect following surgical treatments in patients with UVFP (Morgan et al., 2007; Rosen et al., 2007, 2009; Vinson et al., 2010).

However, the qualitative evaluation of videostroboscopy findings is largely dependent on the experience of the examiner and that in turn may affect the inter- and intra-rater reliability. Objective assessments of the findings have been reported but need sophisticated software and considerable time. Factors that may influence the ability to evaluate the stroboscopic findings are: 1) the evaluator's knowledge on how vocal vibration relates to sound production; 2) knowledge of the anatomy and physiology of larynx; 3) skill with the stroboscopic technique; and 4) skill in interpretation of the recorded images (Hirano and Bless, 1993). A satisfactory level of inter- and intra-rater reliability in evaluation is important to interpret the results. Research tools have been developed to improve the reliability of videostroboscopy evaluation (Poburka, 1999; Poburka and Bless, 1998; Rosen, 2005).

In an attempt to use videostroboscopy as an outcome measure, the inter- and intra-rater reliability has to be determined to inform the reliability of the results of the study.

1.12.2 Visual findings of vocal fold mobility

Objective measurement of airway of glottal area is important in laryngeal research to measure and determine whether an intervention improves vocal fold mobility. Impaired vocal fold mobility leads to a reduction of the glottal airway that causes breathing difficulty (Dailey et al., 2005; Marina et al., 2011b; Waters et al., 1996). To date there is lack of non-invasive methods to objectively measure the vocal fold mobility.

In BVFP, impaired vocal fold abduction during respiration reduces the size of glottal airway hence compromising breathing. Quantification of vocal fold mobility in terms

of the ability of the vocal folds to abduct is necessary to provide a measure of effectiveness of interventions conducted to increase glottal airway area restricted by mobility disorders (Dailey et al., 2005). However, validated, non-invasive quantitative measures are lacking. Glottal area may be assessed during respiration by fibre-optic endoscope commonly attached to stroboscope – videostroboscopy, can be used to quantify vocal fold abduction (Dailey et al., 2005; Motoyoshi et al., 2004; Waters et al., 1996), but accuracy is limited by distortion caused by varying distance from the object and the wide angle lens of the endoscope (Czaja Pawel et al., 2007; Dailey et al., 2005; Forkert et al., 1996). To correct for some of this, it has been proposed to insert a probe to measure the distance from the tip of scope to the true vocal folds, but this increases the size of endoscope required and is too cumbersome, time-consuming and uncomfortable for routine clinical application (Brancatisano et al., 1983). Previously an endoscopic imaging solution has been suggested (Dailey et al., 2005) but valid analysis only appears possible in high quality studies. Specifically, the camera has to be tangential to the plane of the VFs so the images are central position with clear visualisation of anterior commissure and vocal process.

MRI enables high-speed imaging that enables capture of dynamic events non-invasively with good temporal resolution. Such technology, using 0.5-3.0T machines, has been used to capture cardiac, bowel and oesophageal motility patterns (Davis et al., 1996). It has been shown able to capture vocal fold images during adduction and abduction (Ahmad et al., 2006, 2009; Faust et al., 2001; Schlamann et al., 2009). Feasibility of this technology for the detection of vocal fold mobility disorders has been reported (Schlamann et al., 2009). However, the ability to

measure glottal area change from adduction to abduction has not been explored to date.

1.13 Thesis outline

The unifying theme of this thesis is the performance of a series of research studies that collectively amount to a feasibility study for clinical trials of laryngeal reinnervation for the treatment of vocal fold paralysis. However, the results have wider implications for laryngology as will be discussed.

1.13.1 Laryngeal reinnervation in vocal fold paralysis

Laryngeal reinnervation has two types: non-selective reinnervation and selective reinnervation. Non-selective reinnervation is mainly for UVFP to gain normal to near normal voice by re-establishing tone and bulk of the paralysed muscles. Selective reinnervation aims for re-establishment of vocal fold mobility in BVFP to improve glottal airway as well preserving a good voice.

The question as to which treatment is better for UVFP remains outstanding: 'non-selective reinnervation or thyroplasty?', a question that ideally requires an adequately powered and designed RCT. However, RCTs in surgery face inherent recruitment issues related to both surgeons' and patients' views and beliefs that may lead to failure of the trial. These issues and others were explored by carrying out the studies below that will be discussed in detail in **chapter 2.**

I. A national survey of UK ENT consultants exploring their perception towards the trial as well as obtaining crude numbers of eligible UVFP patients under their care for such trial.
II. A qualitative research study employing individual interview of eligible UVFP patients to explore issues around recruitment process as well as to investigate their main voice concern.

1.13.2 Outcome measures to monitor the treatment effect

Treatment outcomes in voice surgery trials are recommended to be multidimensional and standardised. The multidimensional treatment outcomes include subjective as well as objective evaluation. However there are some issues around the objective and subjective evaluations that need investigation.

I. Acoustic analysis is a computer based objective voice signal analysis that bounds to reliability issues due to several confounding effects. A potential software package, OperaVOX, may minimise these issues. However this software has not been rigorously investigated the validity and reliability. Therefore, its performance was compared with MDVP, a commercial software package that widely used in voice research. The results will be discussed in **chapter 3**.

II. Videostroboscopy allows visual evaluation of vocal folds vibratory pattern and the glottal closure. The evaluation is typically subjective with variable inter- and intra-rater reliability. The inter- and intra-rater reliability need to be investigated and reported to inform the strength of the future trial results. Results of this study will be discussed in **chapter 4**.

1.13.3 Non-invasive outcome measures for laryngeal reinnervation

I. LEMG is important in 'ruling in' vocal fold paralysis before reinnervation procedure. In non-selective reinnervation, apart from voice clinical test results, eligibility for the procedure is based on confirmation of paralysis and poor prognosis for vocal fold mobility recovery. LEMG is also the only way presently to show evidence of reinnervation following the procedure as histology evidence is lacking in human studies. However, LEMG is invasive,

operator dependent and not routinely practiced by laryngologists. MRI is a non-invasive tool that maybe an alternative or adjunct to the LEMG. Therefore, a study investigating the feasibility, reliability and validity of T2-MRI to depict TA and PCA muscle bulk and signal intensity in patients with UVFP was done. Results of the T2-MRI study will be discussed in **chapter 5**. The utility of DCE-MRI and DWI in showing perfusion changes of denervated TA and PCA that potentially resolved following reinnervation was explored. Results of the DCE-MRI and DWI study will be discussed in **chapter 6.**

II. Re-establishment of vocal fold mobility in BVFP following selective reinnervation is the main aim. Quantification of vocal fold mobility in terms of the ability of the vocal folds to abduct is necessary to provide a measure of effectiveness of interventions conducted to increase glottal airway area restricted by mobility disorders. Glottal area assessed during respiration by video-laryngostroboscopy can be used to quantify vocal fold abduction, but accuracy is limited by distortion caused by varying distance from the object and the wide-angle lens of the endoscope. Validated, non-invasive quantitative measures are lacking. MRI has the ability to capture vocal fold adduction and abduction. In this thesis, a study aimed to develop and evaluate 3.0T MRI based quantitative measures of vocal fold mobility in normal volunteers and vocal fold paralysis patients will be presented in **chapter 7.**

Development of MRI sequences protocol and evaluation of its feasibility results will be included in **Appendix 6**.

1.13.4 Case series of laryngeal reinnervation

In this thesis, as further preparatory work for a potential randomised controlled trial, I present two small case series:

I. Case series of non-selective laryngeal reinnervation for UVFP will be presented to provide some preliminary results of the surgery before an RCT pilot study is launched. In these case series, multidimensional evaluation of voice outcomes will be presented including the LEMG and T2-MRI results in **chapter 8.1.**

II. Case series of selective reinnervation in rehabilitating unilateral vagal paralysis following vagal nerve tumour resection will be discussed **chapter 8.2**. Selective reinnervation may be the treatment of choice in this condition due to the multicentricity nature in 33% of patients especially those with positive family history (Fang et al., 2011; Urquhart et al., 1994). A successful selective reinnervation that aim at re-establishment of vocal fold mobility as well as improvement of voice may save those patients from irradiation or tracheostomy treatment options if they had a second tumour on the opposite side.

Chapter 2 A qualitative study of the views and beliefs of surgeons and patients on a randomised control trial of laryngeal reinnervation versus thyroplasty for unilateral vocal fold paralysis.

(The amended part on patients' views and beliefs in this chapter has been accepted by the Journal of laryngology and Otology awaiting for publication)

2.1 Abstract

Objectives: A randomised control trial (RCT) of non-selective laryngeal reinnervation versus thyroplasty in unilateral vocal fold paralysis (UVFP) patients was proposed and the present study was to: 1) explore willingness of UK ENT surgeons to recruit patients in to the proposed RCT; and 2) explore UVFP patients' perception of the proposed RCT and their voice concerns.

Methods: A questionnaire was created using SurveyMonkey®. The survey was edited and approved by the ENT UK Survey Guardian prior to distribution. It was circulated via email invitation to the Members of ENT UK, the British Association of Otolaryngologists Head and Neck Surgeons between 21st October and 22nd November 2011. The data of the survey was managed in a spreadsheet and presented in absolute numbers and proportion. Seventeen eligible patients for the RCT, were identified from five voice clinics in London. 11 of 17 patients (18 to 65 years old, 9 females, 2 males), were interviewed using a semi-structured topic guide after a minimum of two weeks to understand the proposed RCT information sheet. The interview was recorded, transcribed and analysed using thematic analysis.

Results: A total of 1194 emails invitation were sent, 318 responded. 296 surgeons see UVFP patients in their clinical practice. 237 surgeons saw a total of 963 per year adult permanent UVFP patients who failed voice therapy. 40% prefer injection laryngoplasty and thyroplasty in managing these patients. 131 responders said they would be willing to participate in the proposed RCT. 100 confirmed by giving their names. 90% of potential trial surgeons are willing to be trained for reinnervation. Forty percent of the eligible patients were willing to be randomised**.** The interviewed

patients were satisfied with the clarity of the proposed RCT information sheet. Most of them perceived that reinnervation was more ‘attractive’ than thyroplasty. This may be due to certain phraseology used in the information sheet and by recruiters. Their main concern was reduced voice strength and the effects of this on work and social life.

Conclusion: The present study of ENT UK survey and qualitative research methods (individual interview) of patients with UVFP, showed that an RCT of reinnervation versus thyroplasty in patients with UVFP is feasible in the UK. The results also highlight potential problem areas for recruitment of surgeon-investigators and patients to this trial. A potential pool of patients for the trial appears promising. The data of individual interview of patients presented evaluates patient-focus and optimisation of the trial protocol, and both recruitment and consent processes. I found phraseology that needed changing or avoiding during the recruitment process. This may in turn improve willingness of potential patients to be randomised. I also proposed VHI-10 to be used as the primary outcome measure in the proposed RCT.

2.2 Background

Patients with unilateral vocal fold paralysis (UVFP) typically experience a hoarse, weak and easily tired voice. Quality of life is significantly reduced and social and employment opportunities are compromised (Fang et al., 2008; Smith et al., 1995). Therefore, maximal rehabilitation of the paralysed larynx is important for health and wealth alike.

Various surgical treatment options have been proposed for UVFP but, to date, there is insufficient evidence concerning the best surgical option to improve voice quality in patients with UVFP. Isshiki type I thyroplasty (Isshiki et al., 1975) is widely used to rehabilitate the permanently paralysed larynx (Leder and Sasaki, 1994; McCulloch and Hoffman, 1998; McLean-Muse et al., 2000; Ryu et al., 2012). It is an operation in which a silastic implant is inserted through a window in the thyroid lamina cartilage of the larynx. This implant medialises the paralysed vocal fold to the midline position allowing the normal opposite vocal fold to make firm contact and produce a stronger voice.

Another approach is non-selective laryngeal reinnervation that restores tone and bulk of the paralysed vocal fold (Crumley et al., 1988) thus re-enables pitch control resulting in normal or near normal voice (Lee et al., 2007; Lorenz et al., 2008; Wang et al., 2011). The innervation to the larynx is re-established by anastomosing the injured RLN and donor nerve. A recent systematic review of studies on reinnervation revealed that there is a need for a prospective trial using standardised, internationally accepted outcomes (Aynehchi et al., 2010). Such a trial, comparing reinnervation and thyroplasty in adult UVFP patients, was attempted in the US (Paniello et al., 2011),

but was suspended prematurely due to issues in obtaining informed consent and low accrual. The investigators found only 24 patients had data suitable for analysis, from an original target population of 298. Within this small group they found no significant differences in voice quality between the two treatment groups. Unfortunately, although providing valuable pointers, the poor governance and data quality of the US trial obviates its usefulness in generating the feasibility data, including for power analysis, that is required for a UK trial sufficiently robust to deliver the information required by patients. This experience highlights some of the commoner problems in surgical trials. The results of this study emphasises the importance of running a feasibility study before a full surgical trial.

Which surgical option delivers superior outcomes for patients with UVFP remains unanswered due to lack of strong evidence in the literature. Therefore replicating the randomised control trial (RCT) comparing non-selective laryngeal reinnervation and thyroplasty is necessary. Nevertheless there remain important uncertainties to be resolved before a new full trial can be initiated. The perception of UK ENT surgeons on surgical techniques in UVFP and on randomisation of their patients specifically, as well as those of the patients themselves, need to be explored. So does the need for training both for reinnervation, an operation not widely practiced presently in the UK, and for thyroplasty (present standard) as the technique may vary from surgeon to surgeon.

Issues in recruitment of eligible patients in such RCT require exploration. It is a problem inherent in RCT in surgery due to factors such as patients' preferences, additional demand of the trial, worry of uncertainty and concern about information

and consent (Ross et al., 1999). It was recommended that patient recruitment is properly planned and piloted hence a feasibility study that includes qualitative methods is necessary (Lovato et al., 1997).

For all these reasons, a feasibility study is a critical step prior to designing a multi-centre, randomised clinical trial of treatment for this disabling condition. Therefore, the primary aims of the present study are:

A. to explore the perception of UK surgeons on surgical techniques in UVFP and on randomisation of their patients in the proposed trial; and
B. to explore the perception of symptomatic UVFP patients concerning the design of an RCT of non-selective laryngeal reinnervation versus thyroplasty.

The secondary aims of the present study are as the followings:

A. Information relating to UK ENT surgeons:-

1. estimate the number of new cases of adults with UVFP seen by UK ENT surgeons annually;
2. establish their current preferences in the treatment of adults with UVFP;
3. investigate which specialists are most willing and able to participate in a full trial;
4. investigate the level of training needed to perform the operations and how best to organise this training;
5. investigate the availability of tests for outcome measures at the ENT surgeons centre; and

B. Information relating to patients:-

1. explore adult UVFP patients voice problems;

2. establish adult UVFP patients' preferences for choice of treatment;
3. what constitutes optimised patient information; and
4. investigate their willingness to be randomised.

2.3 Methods

2.3.1 For UK ENT surgeons (survey)

Study design

This study was carried out by running a survey through questionnaire among UK ENT surgeons.

Questionnaire

The questionnaire was developed through discussions with Laryngologists (MAB, GS, JR, KG) and a triallist (AS). It was created using SurveyMonkey®. The developed questionnaire was then distributed to 10 ENT surgeons for pilot test. Corrections and amendments were done after receiving feedbacks from the ENT surgeons. The final questionnaire was further edited and approved by the ENT UK Survey Guardian (JWF) prior to distribution through the SurveyMonkey website (Appendix 1). Those who were unwilling to participate were invited to give their reasons and comments. Analysis and themes extraction were done whereas for other results, descriptive analysis were made.

Survey – questionnaire distribution

It was circulated via email invitation to the Members of ENT UK, the British Association of Otolaryngologists Head and Neck Surgeons between 21st October and 22nd November 2011.

2.3.2 For UVFP patients (individual interview)

Study design

This study that was approved by the local ethics committee is a prospective qualitative study that involved individual interview.

Participation selection

The study population included UVFP patients who were eligible for the proposed RCT of laryngeal reinnervation versus thyroplasty presenting to voice clinics of five London hospitals in the year 2011-2012. They were then given an information sheet. The principal investigator at the main site, Royal National Throat Nose Ear Hospital, evaluated the patients further within a month to confirm their eligibility. Subsequently consents were sought for the interview session. During the recruitment period, 69 UVFP patients were identified and 17 patients were eligible for the proposed RCT. Of the 17, ten patients (60%) agreed to participate in the RCT and 40% (4 of 10) of them were willing to be randomised (Table 2-1). Eleven patients agreed to be interviewed, 9 females and 2 males with an age range of 18 to 65 years old. Nine of these were English speakers of various ethnicities. Their education level ranged from school leaver to degree holder. The causes of the UVFP were thyroidectomy (five patients), thymectomy (one patient), vagal schwannoma excision (one patient), cervical spine procedure (one patient) and idiopathic origin (three patients). Duration of palsy was between 5 months to 24 months. Of the 11 patients who were interviewed, three patients declined surgical intervention.

Table 2-1: Unilateral vocal fold paralysis patients identified in the clinics

UVFP patients identified – 69			
Not eligible - 52		Eligible – 17	
Reasons	Number	Participate (10)	Not participate (7) (reasons)
Age more than 70	12	4 UVFP patients were willing for randomisation	• Refuse surgical intervention (3) • Only wants IL (2) • Going back to hometown in the United States for LR (1) • Financial issues and not reliable to follow up (1) • Not willing to do extra tests and travelling (1)
Duration of palsy more than 3 years	10		
Multiple lower cranial nerve paralysis	7		
Tumour	4		
Severe lungs problem	1		
Good vocal fold compensation	6		
Prior permanent surgical intervention	8		
Dysarthria	1		
Non-neurogenic	1		
Neuromuscular	2		

LR: laryngeal reinnervation
IL: Injection laryngoplasty

Information sheet

The information sheet explained the trial and included the study design, randomisation, thyroplasty procedure, reinnervation procedure, outcome measures and details of clinic visits.

Interview session – individual interview

The interview was carried out in a dedicated room using a topic guide which was a semi-structured questionnaire developed and based on discussions with ear nose and throat (ENT) consultants and UVFP patient representatives. It focused on the following topics:

1. the patients' voice problems following UVFP,
2. problems encountered due to the voice problems,
3. their understanding of the surgical procedures
4. their understanding of the randomisation
5. what makes them agree or disagree to take part in a research trial
6. what makes them agree or disagree to accept randomisation as part of this trial

The interview was done by a PhD student (myself) who is a trained ENT surgeon. It was recorded digitally and saved as audio files.

Analysis

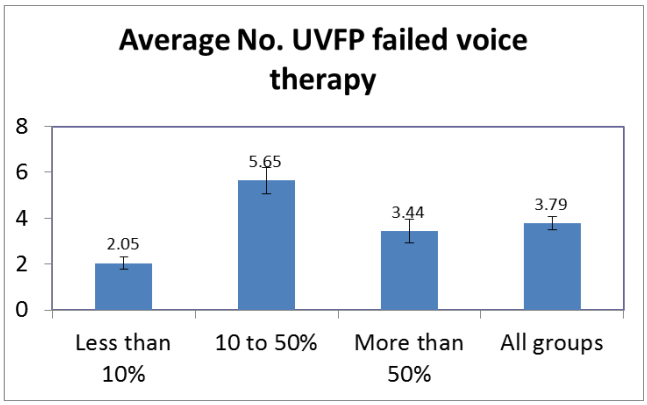
Audio files of the interview were anonymised and transcribed. The transcribed data was analysed using thematic analysis that involved searching across the data to find repeated patterns of meaning (Braun and Clarke, 2006). Specifically, the transcribed text was first familiarised by reading it repeatedly in an active way- searching for

meanings, patterns and issues of potential interest. Pen and paper were used to jot down ideas and extract codes that identify interesting feature of the data. The coding process involved organising the data into meaningful groups. The codes were then collated and potential themes were identified and named. The coding process and themes identification was managed using Microsoft excel.

2.4 Results

2.4.1 For ENT surgeons (survey)

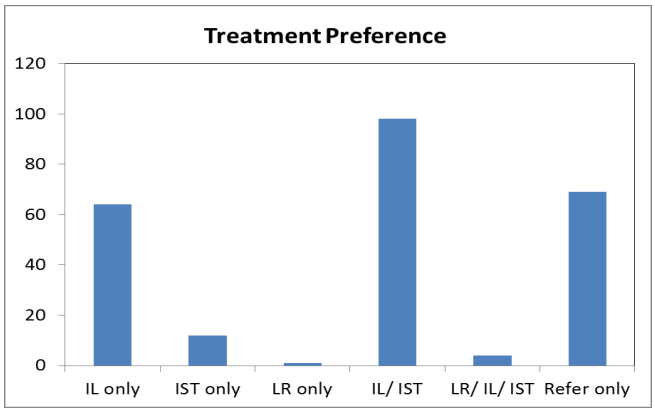
Total of 1194 email invitations were sent and 318 had responded. Of 318 responders, 296 do see UVFP patients in their clinical practice. ENT surgeons who manage adult laryngology patients 'more than 50%' of their patients in clinics is 23.9%, followed by 36.8% and 39.8% for '10 to 50%' and 'less than 10%' respectively. 237 surgeons saw a total of 963 per year adult permanent UVFP patients who failed voice therapy (**Figure 2-1**).



Surgeons were grouped according to percentage of laryngology patients they see in a year as shown in the graph. On average, a surgeon sees 3 UVFP patients who needs surgical intervention in a year

Figure 2-1: Number of UVFP patients failed voice therapy

Forty percent (40%) prefer injection laryngoplasty and thyroplasty in managing these patients. Others choose to perform 'injection only' (26%), 'thyroplasty only' (4%), 'reinnervation only' (0.4%), 'reinnervation with or without medialisation' (1.6%) and 'to refer to other surgeons' (28%) (**Figure 2-2**). General anaesthesia is the main choice for IL (82.6%) whereas for thyroplasty, local anaesthesia is more preferred (68.4%). The most popular implant used for thyroplasty is silastic block implant (57.9%) followed by Gore-tex® implant (35.7%), Montgomery® Implant (10%) and titanium implant (5%). Of 248 surgeons, 9 had performed laryngeal reinnervation in the last 5 years.



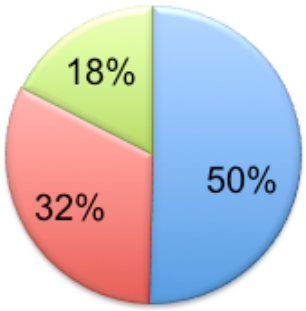
This figure depicts surgeons' treatment preferences. 40% of surgeons who are willing to participate prefer IL/IST in managing patients with UVFP who need surgical intervention

Figure 2-2: Surgeons' treatment preference

IL: Injection laryngoplasty
IST: Isshiki thyroplasty
LR: laryngeal reinnervation

A randomised controlled trial of non-selective laryngeal reinnervation versus thyroplasty was proposed and encouragingly 131 responders said they would be willing to participate in the proposed RCT. One hundred (100) confirmed by giving their names. Their preferred way of trial participation is summarised in (Figure 2-3). Those who are not willing to participate (111 responders) had given their comments. The comments are scrutinised in which themes are extracted and summarised in (Table 2-2).

Regarding surveys on facilities available in centres to measure outcomes, tests that not widely available were laryngeal electromyography (23%) and acoustic studies (40%). Videolaryngostroboscopy, fiberoptic endoscopic examination of swallowing and videofluoroscopy are available in more than 80% of responders' centre and the availability of MRI is 100%.



This figure shows surgeons' preferred way of trial participation which 18% of them prefer to perform both reinnervation and thyroplasty in their centres, 50% prefer to refer to the principle investigator (PI) for both operations and 32% wish to refer to PI for reinnervation and perform thyroplasty in their centres.

Figure 2-3: Surgeons' preferred way of trial participation

Table 2-2: Themes extraction of comments given by surgeons who are unwilling to participate

296 responders, 40% felt unwilling or unable to participate, of which only 6 gave negative comments on reinnervation whereas others give reasons as in the table below.		
Themes	**No**	**Examples of comments**
Negative towards reinnervation	**6**	1. No evidence it could possibly work. Better off spending your time on something more feasible! Sorry!! 2. Very little evidence that it works, except for Mr Tucker himself. 3. I have not seen evidence that laryngeal innervation gives useful functional outcome 4. Thyroplasty works well and is very quick. "If it ain't broke don't fix it" 5. Injections and thyroplasty work and results are immediate. 6. I need to be more familiar with this procedure from my understanding before I would refer a patient directly for this. I would more likely refer a patient to a trusted laryngology colleague who would most likely attempt conventional techniques first. Also, I am not sure if this is ethical, we have one management which we know works and once one is experienced complications and morbidity is low compared to another procedure with multiple techniques which maybe associated with higher morbidity depending on the type of procedure.
Not managing vocal fold palsy primarily	**63**	1. I am a trainee 2. I am an otologist 3. I only see paediatric cases only 4. I would refer to our voice surgeon and the decision would be his, not mine 5. Patients are usually referred to Tertiary referral centre at University Hospital of Wales, Cardiff **6.** Will be retiring in 15 months. Would be happy to refer to colleague if participating
Logistics reasons	**14**	1. I see too few patients to make it worth getting involved 2. I work exclusively in private practice and I think persuading patients to enter a trial would be difficult 3. I wouldn't be able to provide this surgery locally and I doubt my patients would be willing to travel - we are in Cornwall 4. It would depend on whether local centre was involved in trial as many patients do find it difficult/expensive to have to travel to different hospital for treatment and follow-up 5. Good idea but too much other stuff on at present!

Themes	No	Examples of comments
No skill in laryngeal reinnervation	**9**	1. Never done this operation 2. Need to know more about innervation procedure, trial and would need a local operating surgeon 3. I do not think that I would be able to perform laryngeal reinnervation to a good enough standard to propose it to patients as a possible treatment option
Laryngeal reinnervation not suitable for certain group of patients	**6**	1. In patients with UVFP secondary to malignancy (the majority of my cases) I do not consider performing an Isshiki and would always prefer the very quick and less invasive Bioplastique injection. 2. The majority of those I see are paliative patients and benefit from injection, not one of your groups. 3. I am a skull base surgeon, the majority of my patients with unilateral cord palsy have other more significant pathology the management of which takes precedence. In view of this I doubt it would be appropriate for many of my patients to go through this trial. A unilateral cord palsy is usually the least of their worries, if treatment is required the simplest option is favoured. Please define the term 'isolated' above - do you mean no other cranial neuropathy or no other pathology? Were I to see an idiopathic isolated vocal cord palsy I would consider their inclusion, but that's not the group I see. Patients with bilateral vagal palsy would have a lot more to gain from reinnervation, and I would consider referring them for this procedure if it were available (and proven to work).
Not keen on other techniques than injection laryngoplasty	**4**	1. I only do laryngeal injection with Vox (bioplastic) 2. Don't do thyroplasty
No reasons	**9**	

2.4.2 For UVFP patients (individual interview)

Patients' voice problems experience following the UVFP

Voice problems reported by the UVFP patients during the interview are listed in (Table 2-3). The main complaint was that they lost the strength and volume of their voice that affected their work and social life. They also complained of their voice being unpredictable, hoarse and tiring easily. Apart from voice issues, they also had reduced effort tolerance and mild aspiration.

Table 2-3: Voice problems related to unilateral vocal fold paralysis

Voice complains				
Unpredictable	Hoarse	Run out of air	Loss strength and volume	Aspiration
• My voice is very unpredictable • Sometimes my voice goes terribly high • My voice wavers • Someone was to be upset and be agitated, and to be on the verge of tears, that this is how I think they sound like	• Little croaky • Husky	• Walking and talking at the same time very difficult because I ran out of breath • I ran out of breath to finish quick long sentence • Weakness, the fatigue, the volume and strains • My exercise tolerance rate isn't as great • Very difficult and tiring to actually speak to get the volume of air • Gasping for air constantly • Feel out of breath • If I talk without interruption for a longer period, then sometimes I feel dizzy as well • When I speak for a while just get short of breath	• I'm in a place with loud noise I feel like I'm battling a lot • Where there is a lot of ambient noise, I drain out, can't be heard • The communication issues that is the most frustrating in a normal environment like a café or restaurant, a bus terminal • Where it is noisy then I couldn't raise • It's not being able to get that volume level • No power to my voice • I've got no strength in it • I can't convey my message across the crowd • More difficult to speak in a noisy place • More difficult to speak over the telephone • Talking to people on the phone, they can't hear you prop • Lost strength and volume • Sounds as if I had no confidence	• Aspiration occasionally liquid • Swallowing issues are just that I am more conscious of actively swallowing slowly • Sometimes I just choke and I can't breathe and it's very, very painful

Problems encountered due to the voice problems

Problems encountered in the UVFP patients' daily life are summarised in Table 2-4. The UVFP patients suffered decreased self-confidence at work due to the voice problems. They were unable to project their voice in meetings and on the phone to express their opinion:

'Even if I have an opinion, I kind of hold back' (Participant 4)

'I would do other parts of my job and push another senior to do the phone calls' (Participant 4)

'I do so much of my business over the telephone as well as being in large board room scenarios where people, unless they were sitting next to me, wouldn't be able to hear me. Even somebody sitting opposite me on the desk would have difficulty hearing me' (Participant 9)

Those who work outdoors or in a noisy environment, found it hard to carry out their work and one participant resorted to using a voice amplifier.

'Hard to do with my job especially outdoors to raise the voice above a certain level' (Participant 3)

'***The kitchen is so noisy, the machines are on and things, so to scream or to get on top of your voice, you really need a quiet place' (Participant 6)***

An actor who has stopped performing since the voice problem stated:

'***Haven't been able to work for maybe two years now, quite embarrassing sometimes'. (Participant 1)***

Patients isolated themselves from social gatherings, as they couldn't converse effectively with other people. The more they tried, their neck tensed up and the voice got worse and more tiring.

'Socially, it bothers me if I couldn't go out and everyone's talking in a park when I just can't get involved in their conversation so I felt kind of isolated' (Participant 3)

'I used to sing at church and I really enjoy my music. But with this problem now and as soon as I finished it, I go home. I don't stay back because there's no point seeing like, seeing people singing and you can't sing' (Participant 6).

Table 2-4: Problems encountered following the voice problems

Problems arisen due to voice issues	
At work	**At social occasions**
1. Haven't been able to work for maybe two years now, quite embarrassing sometimes 2. Hard to do with my job especially outdoors to raise the voice above a certain level. I have got to use voice amplifying machines 3. Lost a lot of confidence 4. Not being able to speak on the phone if in case I broke down on the motorway 5. Has forced my decision to retire early 6. Hospitals, doctors, they're thinking I have no confidence in myself 7. I would do other parts of my job and push another senior to do the phone calls 8. I'm embarrassed to make a phone call 9. Even if I have an opinion, I kind of hold back 10. Now, I don't shout, I just clap my hands because my voice was giving me a problem 11. The kitchen is so noisy, the machines are on and things, so to scream or to get on top of your voice, you really need a quiet place	1. Talk to the people who are near as me because it's easier 2. Now, I'm the quiet one (less socialised) 3. Socially awkward 4. Social environment, if it's a pub or a restaurant, it was difficult 5. Socially, it bothers me if I couldn't go out and everyone's talking in a park when I just can't get involved in their conversation so I felt kind of isolated 6. When I'm angry, then my voice just cut 7. I used to sing at church and I really enjoy my music. But with this problem now and as soon as I finished it, I go home. I don't stay back because there's no point seeing like, seeing people singing and you can't sing.

Understanding of the surgical procedures

After reading the information sheet and discussing the surgical procedures with ENT consultants prior to the interview, the patients reported that they understood that thyroplasty is an old type of operation that involves an artificial implant which pushes the paralysed vocal fold and is technically easier than laryngeal reinnervation (Table 2-5). They believed that it is temporary in nature and that it may need revision. For reinnervation, they perceived it as a permanent and 'natural' type of operation giving the paralysed vocal fold muscle a new nerve supply that makes it 'alive again' (Table 2-5). However, the way the reinnervation works was not properly understood. Some of them had thought that it would re-establish the paralysed vocal fold movement.

Table 2-5: Unilateral vocal fold paralysis patients understanding of a) thyroplasty operation and b) laryngeal reinnervation operation

a) Understanding on thyroplasty
1. Artificial implant put in to augment the vocal cord
2. Temporary thing
3. There will be an implant to push the paralysed cord forward
4. Put something that will help you
5. The other option seems easier. Just the insertion of the implant, that Seems something you can have under local anaesthetic
6. When reading the sheet, it seemed to me the implant was the easier solution
7. Well, I have read that you don't necessarily need a general anaesthetic
8. I feel that one of the procedures is an older-type procedure
9. Perhaps that it might need to be redone

b) Laryngeal reinnervation (LR)
1. My brain compels messages down to my vocal cord
2. I don't quite understand how the nerves manage to push it over
3. I don't really understand how the nerve one works
4. Giving a muscle new nerve supply
5. Making it alive again
6. Starts to take on its own muscle tone
7. Take on its own muscle movement
8. Take on its own character
9. Relearns how do you…be used again
10. Its new nerve, it could be like learning to talk again
11. Allow a permanent fix as oppose to I guess when you inject with fat or the implant
12. A section of the nerve taking it out from its original position and implanting it into the muscle
13. Maybe I don't understand that one properly
14. Reinnervation, doing something with the nerves, seems like a more natural procedure to do because you're using the body's own materials rather than putting something in that is essentially a foreign body really
15. Natural
16. I was thinking like for cardiac bypass surgery, when the coronary arteries gets clogged, you take it off and reuse the other veins from the leg so, and then the heart comes back to it's normal state even though not 100%. That's the nearest to natural, so that's where I was understanding that process

Understanding about randomisation

All patients said that the information sheet was clear and easily understood. The understanding was further explored by asking them to describe the procedures and randomisation procedure. Eight of 11 UVFP patients understood the concept of randomisation (Table 2-6) and three of them agreed that randomisation was necessary. Only one patient did not really understand the meaning of randomisation although they were still willing for randomisation after a proper explanation had been given.

'***No guarantees as to which you group you will be in' (Participant 2)***

'You get 50-50 chance…. I think for out of 12 people, 6 people would have the injection, 6 people would have the rewiring…But you don't have the choice. You would be…we would be chosen on an ad hoc basis…' (Participant 7)

'Randomisation means that you are randomly selected for one or another treatment. So, it's the one…somebody will decide on your behalf as to how they're going to proceed with the treatment for you or how are they going to deal with it' (Participant 9)

Table 2-6: Unilateral vocal fold paralysis patients understanding on the concept of randomization

Understanding about randomisation
1. Don't know which operation they're having
2. Cut any prejudice
3. Less biased
4. No guarantees as to which you group you will be in
5. You could end up with one or the other
6. You get 50-50 chance
7. I think for out of 12 people, 6 people would have the injection, 6 people would have the rewiring
8. But you don't have the choice. You would be…we would be chosen on an ad hoc basis…
9. Patients don't know which they've got.
10. Randomisation means that you are randomly selected for one or another treatment. So, it's the one…somebody will decide on your behalf as to how they're going to proceed with the treatment for you or how are they going to deal with it
11. So you get put into groups those people will be put into two different groups
12. You shouldn't be able to vote

Agree or disagree for participation in the trial

Those UVFP patients who agreed to participate were motivated by the desire to help other patients, as they believed the outcome of the research would improve the service and offer other patients a better treatment (altruism). Some patients felt that by participating they would get quicker and better treatment as well as better treatment progress monitoring (self-interest) (Table 2-7).

'***I am aware of studies needing to take place to allow progressive movement in research and therefore, allow other people who may have my problem in the future get the right and the best quality of care and treatment for their condition' (Participant 2)***

Conversely, patients who declined to participate were worried about any further surgical interventions (anxiety). They did not want to be offered any surgical options as they felt surgery was invasive and might cause undesired complications. Although not offered by patients themselves as reason for refusing to take part, it is important to note that many had already had a bad experience of surgery causing their UVFP in the first place. Some patients said they were not ready to participate (personal) and that may be related to inability to comply with the extra travel for tests and clinic visits due to busy schedules.

'***I understand the research, but at the moment, I don't think I want to go through it…..too much going on in my life' (Participant 5)***

'I don't even like going to the dentist' (Participant 7)

Table 2-7: What make them agree or disagree to participate

Agree	Disagree
Altruism	**Anxiety**
• I would like to answers to which there aren't any answers to now • I am aware of studies needing to take place to allow progressive movement in research and therefore, allow other people who may have my problem in the future get the right and the best quality of care and treatment for their condition • It will help people • Just something that other people can benefit from the treatment that I will go through • Helps improve the service • My own experience is part of the process	• I understand the research, but at the moment, I don't think I want to go through it • Prefer something less invasive • There could be problems with surgeries • Sometimes it's not 100% effective as we planned • Sometimes there could be complications • I don't even like going to the dentist • I just hate any kind of medical • I really can't face it • If I did it, I would want the easier option
Self benefit	**Personal**
• I think I will get some action done on my voice soon • Active treatment • Put myself forward to have something done instead of sitting. • If your voice doesn't come, then it means you are out of everything, to look for a job, it should be a problem, if you are in the job already, to communicate with people, it's very difficult • And also the information I get back from doing this is a lot puts my mind at ease about the operation at the moment. It just helps because you get more messages to know more about what's going on • Yeah, the treatment will be good for me	• I might in the future, but at present, I feel that I'm not ready • Too much going on in my life

Agree or disagree for randomisation

UVFP patients who agreed for randomisation perceived that both surgical options – reinnervation and thyroplasty - improved voice although one of them was more interested in reinnervation than thyroplasty.

'***I was more intrigued by the nerve one was because in my head, I had the feeling that that would give me back more control however both have the same outcome*****' *(Participant 1)***

'As you'd explained, both of them, all they do is put your vocal cord in the middle and set it so that the other cord makes it easier and you lose less breathe. So they both have the same outcome, you know' (Participant 1)

'Because you're telling me both operations work' (Participant 4)

'Both are good, isn't it?' (Participant 6)

The other factor that encouraged them to agree on randomisation was that they still could have the other surgical option should the first one be unsuccessful.

'***Whichever group I end, it's not contraindicated to having further treatment in the future…. they do both work and it's quicker than waiting…. Because you're telling me both operations work…. whichever operation I have, my voice will improve' (Participant 4)***

Those who did not agree to randomisation had strong treatment preferences. All of the patients who did not want to be randomised chose reinnervation instead of thyroplasty because they believed that reinnervation was permanent and natural whereas thyroplasty was temporary that involved insertion of 'foreign body' or

'plastic' or 'silicone wedge' that may dislodge during vigorous exercise or athletic activity.

'***Nerve graft would've been a once, one off operation if it's successful…. better long term option for my quality of life… seem to be the best option to start with because we could always go back to the thyroplasty if the nerve graft failed…. the other operation the alternative one may not have been permanent…. Reinnervation more of natural thing rather than having an implant… I think it's a lot to do confidence as well' (Participant 3)***

'***I'd prefer the nerve rewiring because I just don't like the idea of having just this…an extra wedge inside I thought it'd be a bit more…natural…. I'd rather prefer to choose…. I don't want to do the old-fashioned way…. personal preference- mainly based on talking to the speech therapist initialy. He recommended it…. silicone wedge over sort of boosting the nerves. I am still quite athletic…. if I had a silicone wedge, I think if I had a knock or a kick to the head…. Because I do boxing and rugby… it might get dislodged. If it didn't work, I would go with…. The wedge is plan-B to me….. From what I have read' (Participant 10)***

I don't like a plastic going to…I don't like the thyroplasty idea… the side effects…. It's more natural -nerve rewiring. I've been researching with myself as well' (Participant 11)

The decision to have reinnervation instead of thyroplasty was influenced by researching, speaking to other patients who had undergone reinnervation and ENT surgeons or speech therapist who saw them in the clinics.

'***I was not prepared to second…for second best. It was clear to me that the best option and the only option for me would be re-innervation but the other two options are far too temporary. I thoroughly researched all of the options available. I suppose PB (ENT consultant) had given me some fairly clear indication just by being frank that the fact the other tops of medialisation, the thyroplasty, the fact that it wasn't something that at this stage was expected to last longer than about 10 years, I believe MR (another ENT consultant) indicated that it was, you know, that both of those again were more shorter term options' (Participant 9)***

Although they were aware that they still can have reinnervation if the thyroplasty did not work, they worried that they may not be eligible anymore due to cut off duration of palsy for the trial also due to age as reinnervation is thought to be less effective as someone become older.

'***I wouldn't be able to have the re-innervation after three years, after my initial injury, which would be October 13….I am 62 and my voice will probably change as I get older, I suppose I want to feel I have done everything I can to preserve my voice so that it won't actually deteriorate too much' (Participant 8)***

2.5 Discussion

2.5.1 Overview

RCT is widely proposed as the best method to evaluate the efficacy of certain treatments or surgical interventions. However, RCT in surgery encounters methodological and practical challenges greater than those of drug trials due to factors including varying surgeons' and patients' equipoise in which one surgical option is preferred over the other. These challenges can lead to poor recruitment (Abraham et al., 2006; Ergina et al., 2009; McCulloch, 2002; Ross et al., 1999). These authors recommend that a feasibility study employing qualitative research methods is carried out so that patients' recruitment would be properly planned (Lewin et al., 2009; Lovato et al., 1997; Murtagh et al., 2007; Ross et al., 1999). Good qualitative studies deliver a major contribution in the success of RCT in surgery. Therefore I performed this qualitative study before embarking on the proposed RCT to explore surgeons' willingness and identify problems around patients' recruitment also their main voice concerns.

2.5.2 For ENT surgeons (survey)

Failure of any RCTs causes waste of money and time of researchers that also discourage them from carrying out future research. One of major barriers in an RCT of surgical trials is a failure to recruit adequate number of participants as may be only subgroup of population are eligible and some patients may have co-morbidities that precluded them from the surgery (McLeod, 1999). According to this survey, numbers of UVFP patients who are potentially to be recruited are promising. Average of UVFP patients eligible for surgical intervention is 3 patients per surgeon. 137 are willing to participate and 100 had confirmed by giving out their names. So

potentially it is possible to get 300 patients per year. The US trial targeted 298 patients before it was stopped prematurely at 60 patients due to irregularity in obtaining inform consent and low accrual (Paniello et al., 2011). This number potentially could be achieved by multicentre participation in the UK as the figures here demonstrate (McLeod, 1999).

The proposed RCT in the present study is designed to investigate which surgical technique is more effective. Therefore it requires involvement of surgeons of multicentre so the conclusions drawn from the trial has general application (Lilford et al., 2004). However surgeons' willingness to participate in a RCT is one of the barriers that would potentially affect patients' recruitment. Strong surgical preference or lack of expertise in performing a new procedure may preclude a RCT (Solomon and McLeod, 1995). However the strong surgical preference may not reflect the surgeons' willingness to recruit their patients (Litton and Delaney, 2013; Palmer et al., 2013). According to this survey, most of ENT surgeons in the UK prefer injection laryngoplasty and thyroplasty in managing their UVFP patients who failed voice therapy and reinnervation is the least preferred. Although ENT surgeons are not familiar with the reinnervation procedure, they are willing to be trained (90%) and 18% of them wish to perform reinnervation in their own centres. Training on reinnervation is then very important before embarking on an RCT to compare it with a more familiar surgical technique, thyroplasty. The next step after training is standardisation of the operative technique so the participating centres reach a set standard. ENT surgeons who have been trained should perform reinnervation in their own centres and a regular assessment should be made so the set standard is achieved.

Among 111 responders who were unwilling to participate, only 6 are negative towards reinnervation down to their belief that it will not work on patients with UVFP. The rests gave reasons that are not directly related to refusal to let their patients to be randomised. Specifically, they are concerned about their unfamiliarity of the reinnervation procedure, unsuitability of reinnervation for cancer patients with poor prognosis and patients have to make a long travel to get this done. Majority of them are trainees in ENT who said that they would participate if the local consultants decide to participate. So potentially they would be able to recruit participants and send them to the participating consultants.

Another factor that is important before embarking a multicentre RCT is standardisation of outcome measures among centres. This survey asked the availability of facilities to measure outcomes such as acoustic studies, video-laryngostroboscopy, laryngeal electromyography, aerodynamic studies and MRI. Laryngeal electromyography is important in giving evidence of reinnervation although it is not the primary outcome measure. Evaluation of laryngeal muscle characteristics before and after reinnervation is important and adds an understanding of mechanisms of denervation and reinnervation. Reinnervation should not be performed without confirming vocal fold paralysis as in some cases vocal fold immobility maybe due to mechanical fixation. However laryngeal electromyography is only available in 23% percent of centres. MRI could be an alternative as it was found to be available in 100% of centres. With its advanced technique, potentially more information of the paralysed vocal fold could be obtained such as the muscle bulk, fatty degeneration and perfusion of the muscle (Kamath et al., 2007; Koltzenburg and Yousry, 2007).

Measurement of voice quality is important in the proposed trial. There was major diversity in treatment outcome measures with regards to functional voice assessment used in the literature. The major diversity hinders meta-analyses of the results of voice treatment. A basic protocol for pre operative and post operative functional voice assessment had been proposed to achieve uniformity in the voice research methodology and to allow relevant comparisons with the literature when presenting or publishing the results (Dejonckere et al., 2001). The protocol includes voice perceptual analysis by professionals (grade, roughness, breathiness), video-laryngostroboscopy evaluation (closure, regularity, mucosal wave and symmetry), acoustic analysis (jitter, shimmer and F0), aerodynamic analysis and subjective rating by the patient. Unfortunately acoustic studies are only available in 40% of centres and they are of different type of machines.

2.5.3 For UVFP patients (individual interview)

Recruitment process

I explored the patients' views and beliefs after they had read an information sheet about the proposed study. Ethical committees require information sheets to ensure shared treatment decision making between clinicians and patients (Charles et al., 1998). It is recommended that sheets are clear but concise in explaining the study design, details of procedure, advantages and side effects (Ellis et al., 2002; Mills et al., 2003). The sheets can however influence patients' perception and reasoning and ultimately affect the recruitment to a trial (Donovan, 2002; Donovan et al., 2009; Edwards et al., 1999; Mills et al., 2003). Recruiters may use terminology that subconsciously encourages potential patients to prefer one treatment to another (Donovan, 2002; Donovan et al., 2009; Mills et al., 2003). Donovan et al. have shown that the rate of randomisation was boosted (Donovan, 2002; Donovan et al., 2009) after the terminologies were changed in the written informed consent.

In the present study, patients were given a minimum of two weeks to read through the information sheet about the study before the interview. All patients were satisfied with the clarity of the content. For the description of the operations, the patients tended to describe reinnervation as a 'natural' and 'permanent' type of operation that make the vocal fold 'alive' again. Whereas for thyroplasty, they tended to explain it as a 'temporary', 'old-fashioned' type of operation using a 'wedge' or 'plastic' inserted into the neck to make the voice stronger, and an operation that may need revision at some point. Presentation of the options in the literature provided verbally by recruiters led to the patient perception that reinnervation might be more 'attractive' than thyroplasty, a clear source of bias and lack of equipoise. The fact

that the reinnervation cannot presently re-establish normal vocal fold movement, but rather aims to improve tone and bulk of the vocal fold muscles should be conveyed effectively. The information sheets and recruiters also should emphasise that both type of operation have been shown repeatedly to be successful in strengthening UVFP patients' voices, but we simply do not know which one is better at addressing those concerns most important to patients themselves.

Factors that motivated patients to participate in the present study were similar to reasons reported by other studies such as altruism and self benefit (Daugherty et al., 1995; Featherstone and Donovan, 2002). Conversely, patients who refused to participate regarded the operation as invasive, had strong treatment preferences and regarded participation as time consuming. Similar reasons were also mentioned in previous studies (Alderson, 1996; Bevan et al., 1993; Schwartz and Fox, 1995). Eighty per cent of UVFP patients in this study understood about randomisation and the importance of it but could not accept the concept of equipoise or had a strong treatment preference. They wanted to be in control in making decisions as to which surgical option they would receive. UVFP patients who accepted that 'both treatment work to improve the voice'- clinical equipoise- tended to accept randomisation.

Voice concerns

McCulloch et al. suggested surgical researchers run a feasibility study to identify suitable primary outcome measures (McCulloch, 2002). There has been no consensus to date on the primary outcome measures for trials evaluating the efficacy of treatments for UVFP. Each of the three RCTs published to date used different primary outcome measures such as self-reported visual analogue scores of voice quality (Hertegård et al., 2004), the well-validated voice handicap index (VHI) (Jacobson et al., 1997; Lau et al., 2010) and perceptual evaluation of voice by untrained listeners (Paniello et al., 2011). However, crucially, surgeons' choice of outcomes might not truly represent patients' concerns (Ergina et al., 2009). Here I explored UVFP patients' concerns about their voice to determine the primary outcome measure that is most likely to be meaningful to patients in a full trial. It was found that their main concern was, reduced voice strength and volume and the effects of this on work and social life. Therefore, I propose a short and fully validated voice-specific self-reported outcome tool should be used in future trials.

Quality of life outcome measure has been increasingly used in clinical trials (Calvert et al., 2011; Temple et al., 2009). Voice-specific patient self-reported outcome is a disease specific quality of life measures. It has high content validity since the patients have been living with the voice so their satisfaction may reflect success of treatment (Carding et al., 2009). The tool enables assessment of level of handicap and impact of voice problems to the quality of life. This kind of outcome measure also allows assessment of the patient's ability to use his or her voice at working or social environment (Hogikyan and Rosen, 2002). Furthermore objective

measurements of voice on one particular day and time may not represent the patient's overall voice quality. The voice may vary throughout the day or worse at the end of working day or fluctuate from one day to another day. Therefore the voice-specific patient self-reported outcome may be the most appropriate tool to assess the overall voice quality, patients' satisfaction and success of interventions.

However the patient-self reported outcome tool is limited by its subjectivity. Patients' relative concept of quality of life, degree of expectation and cultural differences may affect the scores of outcome. It may also introduce biases to the results of a RCT especially if blinding of patients is not possible. Therefore objective outcome measures need to be included in the trial to confirm or inform the strength of the results.

There are a number of validated voice-specific patient self-reporting tools in the literature. They are used to evaluate the impact of voice disorders to patients' life and daily activity. Voice Handicap Index (VHI), Vocal Performance Questionnaire (VPQ) and Voice Symptom scale (VoiSS) have been used widely in voice research. All of these tools were shown to have excellent internal consistency (Cronbach's coefficient = 0.81 to 0.95) (Carding et al., 2009b). Of the three tools, VHI reached the highest repeatability with intraclass correlation of 0.83. However VHI is a long questionnaire that consists of 30 questions. Rosen et al. had done further factor analysis, shortened the VHI to 10 questions (VHI-10) and validated it (Rosen et al., 2004).

VHI-10 is a questionnaire that incorporates 10 items measuring the ability of voice to compete against background noise and project. It comprises three domains: functional (5 items), physical (3 items) and emotional (2 items). Deary et al. compared VHI-10 and VPQ as both are short and convenient. It was found that both are internally consistent and good overall indicator of the severity of voice disorders (Deary et al., 2004). The VHI-10 has been used in other retrospective or prospective trials involving patients with UVFP and was able to demonstrate treatment effect before and after operation (Carding et al., 2004; Hogikyan and Rosen, 2002; Misono and Merati, 2012; Rosen et al., 2000).

2.5.4 Conclusion

The present study of ENT UK survey and qualitative research methods (individual interview) of patients with UVFP, showed that an RCT of reinnervation versus thyroplasty in patients with UVFP is feasible in the UK. The results also highlight potential problem areas for recruitment of surgeon-investigators and patients to this trial. A potential pool of patients for the trial appears promising. The majority ENT surgeons in the UK who manage patients with UVFP are receptive to the proposed trial and willing to be trained. Trainings on laryngeal reinnervation are vital as not many surgeons are familiar with this surgery. The data of individual interview of patients presented evaluates patient-focus and optimisation of the trial protocol, and both recruitment and consent processes. I found phraseology that needed changing or avoiding during the recruitment process. This may in turn improve willingness of potential patients to be randomised. I also proposed VHI-10 to be used as the primary outcome measure in the proposed RCT.

Chapter 3 Reliability of OperaVOX against Multidimensional Voice Program (MDVP)

(The amended part of this chapter has been accepted by the Clinical Otolaryngology Journal awaiting for publication)

3.1 Abstract

Objective: To evaluate the agreement between OperaVOX and MDVP.

Methods: Fifty healthy volunteers and 50 voice disorder patients had supervised recordings in a quiet room using OperaVOX by the iPod's internal microphone with sampling rate of 45kHz. A five seconds recording of vowel /a/ was used to measure fundamental frequency (F0), jitter, shimmer and noise–to–harmonic ratio (NHR). All healthy volunteers and 21 patients had a second recording. The recorded voices was also analysed using the MDVP. The inter- and intra-software reliability was analysed using intraclass correlation (ICC) test and Bland Altman (BA) method. Mann Whitney test was used to compare the acoustic parameters between healthy volunteers and patients.

Results: Nine of 50 patients had severe aperiodic voice. The ICC was high with a confidence interval of >0.75 for the inter- and intra-software reliability except for the NHR. For the inter-software BA analysis, excluding the severe aperiodic voice datasets, the bias (95% LOA) of F0, jitter, shimmer and NHR was 0.81(11.32, -9.71); -0.13 (1.26, -1.52); -0.52 (1.68, -2.72); and 0.08 (0.27, -0.10). For the intra-software reliability, it was -1.48 (18.43, -21.39); 0.05 (1.31, -1.21); -0.01 (2.87, -2.89); and 0.005 (0.20, -0.18), respectively. Normative data from the healthy volunteers was obtained. There was a significant difference in all acoustic parameters between volunteers and patients measured by the OperaVOX ($p<0.001$) except for F0 in females ($p=0.87$).

Conclusion: OperaVOX is comparable to MDVP and has high internal consistency for measuring the F0, jitter and shimmer of voice except for the NHR.

3.2 Background

There is an increasing requirement for robust measures of voice quality, as a means of assessing outcomes of treatments for voice disorders and airway surgery. These have value to the patient, the individual surgeon, unit audits and research, as well as for risk management. However, establishing such measures in a way that satisfies both clinical and truly patient-facing needs has proved elusive. One reason is that assessments of voice need to be multidimensional. They need to involve perceptual as well as objective evaluation such as acoustic analysis (Dejonckere et al., 2001).

Acoustic analysis is a non-invasive computer based objective assessment of voice quality by measuring the properties of recorded voice signals. The common measurements includes fundamental frequency (F0), jitter, shimmer and noise-to-harmonic ratio (NHR) (Amir et al., 2009). Jitter, shimmer and NHR, are the most common parameters reported in voice research (Brockmann-Bauser and Drinnan, 2011). F0 is a measurement of the mean frequency of mucosal vibrations of the vocal folds. Jitter and shimmer are perturbation measurements that measure cycle-to cycle frequency and amplitude variation respectively in the analysed voice sample. NHR is a measurement of degree of hoarseness obtained by estimating the proportion of noise in the subject's voice.

A comprehensive, systematic and routine measurement of voice following treatment is only possible in hospitals with voice laboratory facilities. There are many software programmes available to measure acoustic parameters. Multidimensional Voice Program (MDVP, KayPentax, NJ, USA) is a commercial software package widely used in research that is available at the present study's recruitment centre. Co-workers at UCL, with commercial partners Oxford Research Wave Ltd., have

developed OperaVOX (On Person RApid VOice eXaminer), portable voice analysis software running on an iPod, iPhone and iPad (Apple, USA), in an attempt to meet the specifications above.

In the present study, the aim was to evaluate the agreement between OperaVOX and MDVP by measuring the inter- and intra-software reliability.

3.3 Methods

Study design

This cross-sectional reliability study has been approved by the Local Ethics Committee (reference number: 11/LO/0583).

Participants selection - One hundred voice samples of adult volunteers (50) and patients (50) with voice disorders presenting to the four laryngologists at the Royal National Throat Nose and Ear Hospital, London, were recruited using convenience sampling and their demographics are summarised in (Table 3-1). The volunteers had a mean age of 35.14 (s.d.10.52) years' old, and no previous voice problems. Those aged more than 65 years' old, smoking history, vocal fold pathologies, history of endotracheal intubation for the past six months, current or the past two weeks history of upper respiratory tract infection, voice handicap index-10 (VHI-10) of more than 11 or reflux symptom index of more than 7 were excluded. The patients had a mean age of 48.98 (s.d. 14.35) years' old and were diagnosed with a wide range of voice disorders, including: muscle tension dysphonia (MTD), laryngopharyngeal reflux (LPR), vocal fold paralysis, spasmodic dysphonia, sulcus vocalis, and vocal fold lesions. In all patients with vocal fold pathologies, diagnoses were supported and documented using a high-resolution digital fibreoptic nasopharyngolaryngoscope system (VNL-1590STi, KayPentax, USA).

Table 3-1: Demographic data. Where means are shown, brackets represent standard deviation.

	Volunteers	Patients
Number	50	50
Mean age	35.14 (10.52)	48.98 (14.35)
Male	22	35
Female	28	15
English first language	40	39
English second language	10	11
Diagnosis		
Muscle tension dysphonia	-	11
Laryngopharyngeal reflux	-	1
Vocal fold paralysis	-	23
Spasmodic dysphonia	-	8
Sulcus vocalis	-	2
Vocal fold lesions	-	5
VHI-10	1.98 (2.41)	23.54 (8.00)
Severe aperiodic voice	-	9
Retest	50	21

Study software - OperaVox™, is a software application that allows voice recordings as well as performs on-device acoustic analysis for users either in the hospital or at their homes. In this study the software was installed on an iPod touch 4th generation (Figure 3-1). A lanyard was attached to it for the purpose of standardisation of 30cm lips-to-device distance (Figure 3-2).

Instructions were provided on the page prior to the test to guide patients on how to perform the test and standardise the procedure. It also has a colour bar indicator to measure instantaneous loudness of the voice while the recording was performed (-6 dBFS is upper end of green and -24 dBFS is lower end of green).

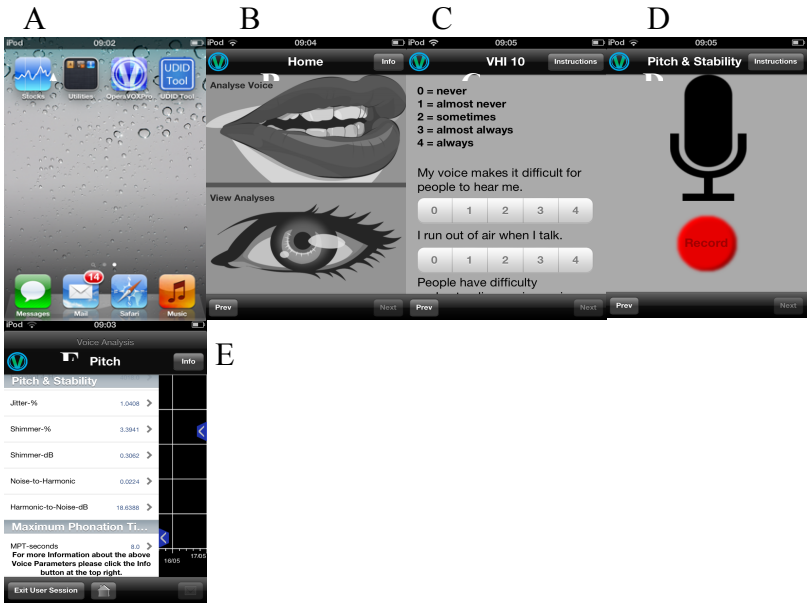


Figure 3-1: This figure shows OperaVOX that has been installed on an iPod touch that shows: A) software icon; B) start analysis page; C) VHI-10 page; D) voice recording page; and E) results page.

Figure 3-2: This photo shows a participant holding the iPod taut during the voice recording

Multidimensional Voice Program is a widely used voice analysis software in published voice research.

Supervised recordings were performed in a quiet room with noise threshold of -39dBFS, using OperaVOX by the ipod's internal microphone with a sampling rate of 45kHz. The start of a recording was disabled in the presence of any noise above a pre-determined threshold value. They were instructed to do the recording while sitting on a chair without any accessories worn on the wrist so as to avoid any background noise from being recorded onto the recorded voice. Holding the iPod taut with a lanyard that was worn around the neck standardised the lips-device distance. They were instructed to say vowel /a/ for 6 seconds at comfortable loudness and to keep the colour indicator within the green range while recording to ensure there is sufficient loudness and thereby signal to noise for accurate measurements.

A five seconds recording of a sustained vowel /a/ was used by OperaVox to measure F0, jitter percent, shimmer percent and NHR. The first 500ms of the start and the end of the vowel was removed to avoid the start and end effect. Once satisfied with the recording, the participant pressed the 'yes' button to perform immediate analysis. A second recording was done within 15 minutes after the first. All voice samples were saved in wav.files and analysis was also performed using MDVP for comparison. Patients whose voices were severely aperiodic (Titze, 1995) were excluded from certain parts of statistical analysis.

Statistical analysis

Inter- and intra-software reliability was assessed by using intraclass correlation coefficient (ICC) (Lee et al., 1989; Shrout and Fleiss, 1979) and analysing the bias and 95% limit of agreement (LOA) of the Bland-Altman (BA) plot (Bland and

Altman, 1999), of parameters produced by OperaVOX and MDVP (inter-software), and between the first and second recording of each software (intra-software). The BA analysis was repeated after excluding severe aperiodic voice datasets. ICC values ranged from 0 (no correlation) to 1 (perfect correlation). The ICC has to be at least 0.6 for the measurements to be clinically useful (Chinn, 1991). Reliability of the measurements was considered good for ICC of similar to 0.75 and moderate for ICC of between 0.5 and 0.75 (Portney and Watkins, 1993). Normative data of volunteers produced by each software programme was presented. Acoustic parameters were compared between volunteers and patients (excluding the severely aperiodic voices) using Mann Whitney test.

3.4 Results

The data presented in (Table 3-2) was not normally distributed therefore median and interquartile range (IQR) was used.

Table 3-2: Comparison between healthy volunteers and patients with voice disorder in females and males after excluding severe aperiodic voice. Where medians are shown, brackets represent interquartile range (IQR).

	Female			Male		
	Volunteers N = 28	Patients N = 28	p value	Volunteers N = 22	Patients N = 13	p value
F_0 M	219.43 (197.05, 235.73)	204.54 (192.92, 247.70)	**0.87**	116.38 (108.57, 122.06)	139.78 114.60, 169.34)	**0.006**
F_0 O	219.76 (197.18, 232.96)	203.85 (193.45, 243.38)	**0.83**	116.68 (108.39, 122.04)	139.83 (114.66, 161.60)	**0.006**
Jitter (%) M	0.71 (0.51, 0.88)	2.31 (1.24, 3.26)	**0.001**	0.60 (0.47, 0.79)	1.87 (1.19, 2.74)	**0.001**
Jitter (%) O	0.69 (0.48, 0.90)	2.85 (1.16, 3.90)	**0.001**	0.54 (0.44, 0.75)	1.83 (1.15, 2.86)	**0.001**
Shimmer M (%)	3.16 (2.63, 4.01)	6.21 (5.06, 8.70)	**0.001**	4.46 (3.43, 5.36)	6.32 (4.88, 8.84)	**0.003**
Shimmer O (%)	3.46 (2.83, 4.47)	7.72 (5.27, 10.33)	**0.001**	4.30 (3.33, 4.92)	8.10 (5.03, 8.89)	**0.001**
NHR M	0.13 (0.12, 0.13)	0.16 (0.13, 0.19)	**0.004**	0.14 (0.14, 0.15)	0.16 (0.14, 0.16)	**0.005**
NHR O	0.01 (0.006, 0.03)	0.06 (0.02, 0.24)	**0.001**	0.03 (0.02, 0.04)	0.05 (0.02, 0.07)	**0.05**

M = MDVP, O = OperaVOX, p = p value of Mann Whitney test

Inter-software reliability

There was high inter-software agreement for F0, jitter and shimmer with ICC (95% CI) of 0.98 (0.97,0.99); 0.94 (0.91, 0.96); and 0.93 (0.88, 0.96) but for NHR the ICC was 0.52 (0.36, 0.65). The bias (95% LOA) of all the parameters before exclusion of the severe aperiodic voices was -0.49 (21.16, -22.14); -0.07 (1.44, -1.58); -0.91 (3.44, -5.26); and 0.01 (0.60, -0.58) respectively. After the exclusion, the bias (95% LOA) was 0.81(11.32, -9.71); -0.13 (1.26, -1.52); -0.52 (1.68, -2.72); and 0.08 (0.27, -0.10) (Figure 3-3).

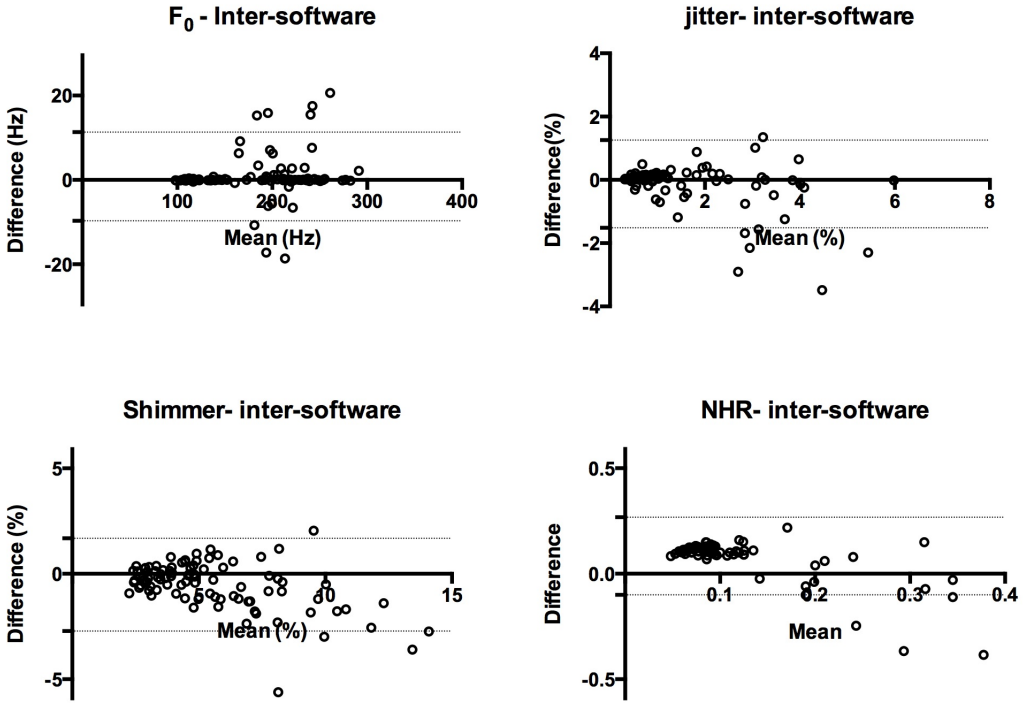


Figure 3-3: Inter-software Bland Altman plot after excluding severe aperiodic voice. Dotted lines represent upper and lower 95% limit of agreement.

Intra-software reliability

There was high intra-software ICC of F0, jitter and shimmer for OperaVOX: 0.93 (0.89, 0.96); 0.92 (0.91, 0.97); and 0.96 (0.94, 0.98), and for *MDVP*: 0.92 (0.88, 0.95); 0.89 (0.83, 0.93); and 0.91 (0.87, 0.94). For NHR, the ICC (95% CI) was 0.64 (0.48, 0.76) for the MDVP, and 0.72 (0.58, 0.81) for the *OperaVOX*.

For the MDVP, the bias (95% LOA) for F0, jitter, shimmer and NHR before exclusion of the severe aperiodic patients were –3.92 (38.73, -46.56); 0.12 (2.08, -1.85); -0.33 (5.03, -4.36); and 0.02 (0.21, -0.21). After the exclusion, they were -0.88 (18.21, -19.96), -0.05 (1.03, -1.12), 0.12 (2.84, -2.61) and 0.002 (0.07, -0.07) (Figure 3-4).

For the OperaVox, the bias (95% LOA) of F0, jitter, shimmer and NHR before exclusion of the severe aperiodic voices were -1.55 (37.14, -40.23); 0.11 (1.53, -1.31); 0.08 (3.93, -3.77); and 0.03 (0.64, -0.58). After the exclusion, they were -1.48 (18.43, -21.39); 0.05 (1.31, -1.21); -0.01 (2.87, -2.89); and 0.005 (0.20, -0.18) (Figure 3-5).

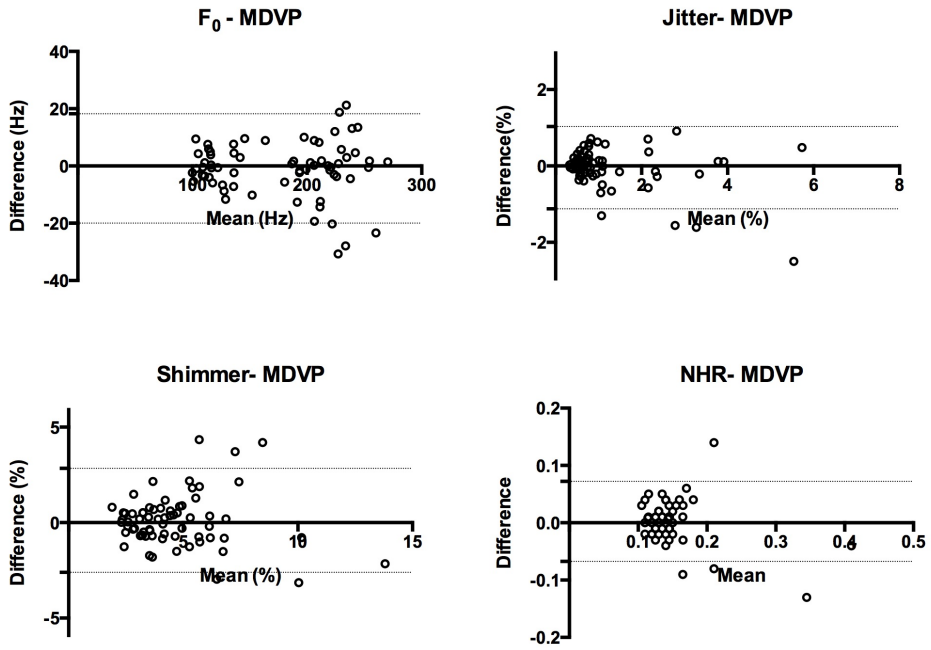


Figure 3-4: Intra-software Bland Altman plot for MDVP after excluding severe aperiodic voice. Dotted lines represent upper and lower 95% limit of agreement.

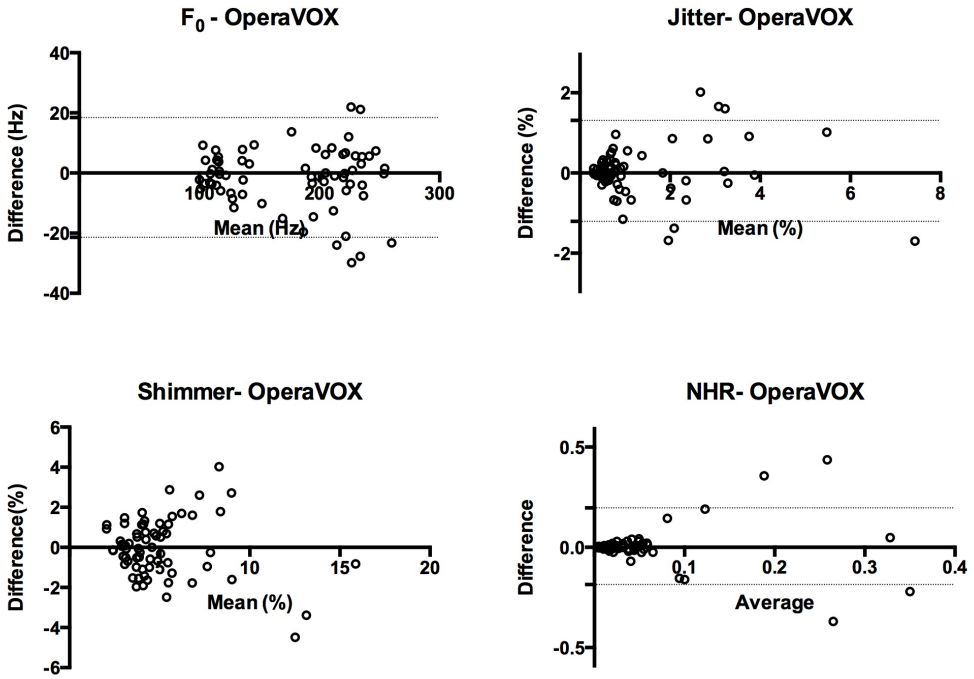


Figure 3-5: Intra-software Bland Altman plot for OperaVOX after excluding severe aperiodic voice. Dotted lines represent upper and lower 95% limit of agreement.

Normative data

Normal range of all the parameters of volunteers from MDVP and OperaVOX is summarised in Table 3-3.

Table 3-3: Normative data measured by OperaVOX and MDVP. Where means are shown, brackets represent standard deviation.

	Female (n = 28)		Male (n =22)	
	MDVP	OperaVOX	MDVP	OperaVOX
F_0 (SD)	217.76 (31.58)	217.85 (30.80)	116.30 (11.58)	116.29 (11.51)
Jitter (SD)%	0.70 (0.25)	0.80 (0.44)	0.71 (0.25)	0.65 (0.23)
Shimmer (SD) %	3.18 (0.90)	3.56 (0.94)	4.43 (1.31)	4.23 (1.26)
NHR	0.13 (0.02)	0.03 (0.08)	0.14 (0.01)	0.03 (0.02)

Comparison between healthy volunteers and patients

There was significant difference in the F0, jitter, shimmer and NHR between volunteers and patients from the two software programmes ($p<0.001$ or $p<0.05$) except for F0 ($p = 0.87$) of female participants from both OperaVOX and MDVP (Table 3-2). The measurements were found higher in the patients group than the volunteers indicating abnormal voice.

3.5 Discussion

Acoustic analysis is useful in treatment monitoring following any interventions for vocal fold pathologies. It measures irregularity of mucosal vibration of the vocal folds that may affect the F0, jitter, shimmer and NHR. The acoustic parameters were documented in papers reporting voice improvement in unilateral vocal fold paralysis patients following surgical intervention (Chowdhury et al., 2013; Hassan et al., 2011; Ryu et al., 2012; Schneider-Stickler et al., 2013; Wang et al., 2011).

Previous studies suggested that inter- and intra-software acoustic parameters were only moderately correlated (Carding et al., 2009), therefore they are not recommended to be used independently (Brockmann-Bauser and Drinnan, 2011). The lack of reliability may be due to several confounding effect such as different algorithms used by different software programmes (Bielamowicz et al., 1996), recording environment, type of microphone, frequency and amplitude of phonation during recording and distance of the microphone from the lips (Brockmann-Bauser and Drinnan, 2011). Some of the confounding effects may have been effectively controlled by the OperaVOX technology as some of its features seek to ensure better standardisation of recording. For instance, voice recordings by the OperaVOX are only made possible at an acceptable level of ambient noise. The built in colour indicator ensures that the participants project their voice during recording so as to conform to accepted levels of loudness. Standardisation of lips-to-device distance can be achieved using a lanyard. Voices recorded by an internal mobile Apple device's microphone was shown to be compatible with a direct digitisation method (Hornibrook et al., 2012; Lin and Hornibrook, 2011).

Strength of the study

OperaVOX reliability has not previously been rigorously investigated, so in this study it was compared with another widely available commercial programme - MDVP in measuring F0, jitter, shimmer and NHR. We employed recommended statistical analysis - ICC and BA analysis to measure inter- and intra-software reliability of the new voice analysis software package.

Synopsis of key findings

Inter-software reliability showed high agreement in all parameters except the NHR. The BA analysis showed that OperaVOX and MDVP measured F0, jitter, shimmer and NHR with bias of -0.49, -0.07, -0.9 and 0.008 respectively. Intra-software reliability was good to excellent for all parameters. The BA analysis demonstrated bias of -1.55, 0.11, 0.008 and 0.003 for the respective parameters. When the severely aperiodic voices were removed, more precision was achieved for the inter- and intra-software BA analysis in which the variation of bias and 95% LOA became narrower. This was clinically acceptable to the authors' opinion. We also demonstrated that acoustic parameters measured by OperaVOX were able to differentiate normal and abnormal voices. These results indicate that OperaVOX is comparable to MDVP. However, data transfer between the two software programmes is not recommended.

Previous studies have shown that normal range measured by different programmes (

Table 3-4) and in different population varies (Dehqan et al., 2010; Felippe et al., 2006; Godino-Llorente et al., 2008; Ma and Yiu, 2005; Toran and Lal, 2010). Normal range for all parameters measured by OperaVOX was close to the MDVP data except the NHR.

Table 3-4: Normative data from previous studies

	Gender	F0, Hz	Jitter, %	Shimmer, %	NHR	HNR
MDVP programme (from the manual)	Female	241.08(25.107)	0.633 (0.351)	1.997 (0.791)	0.112 (0.009)	
	Male	141.74(21.136)	0.589 (0.535)	2.523 (0.997)	0.122 (0.014)	
2009 Toran and Lai (Nepalian)	Female	246.45 (39.73)	0.14 (0.04)	1.57(0.43)	-	25.88(2.75)
	Male	170.05 (32.78)	0.14 (0.02)	1.66(0.47)	-	25.81(2.62)
2010 Dehqan et al. (Iranian)	Female	214.64 (1.16)	0.22 (0.01)	1.21 (0.03)	-	18.81 (0.96)
	Male	112.82 (0.94)	0.23 (0.02)	1.22 (0.02)	-	18.42 (0.57)
2006 Fellippe et al. (Brazilian)	Female	206	0.62	-	-	10.98
	Male	120	0.498	-	-	9.56
2005 Ma and Yiu (Hong Kong)	Female	226.06 (21.46)	-	6.07 (1.17)	0.24 (0.04)	
	Male	139.85 (12.19)	-	8.89 (0.11)	0.31 (0.01)	
2008 Godino-Ilorente MDVP	Female	199.98 (31.4)	0.99 (0.70)	3.06 (1.41)	-	11.6
	Male	120.56 (24.8)	0.96 (0.78)	3.53 (1.78)	-	12.2
2008 Godino-Ilorente WPCVox	Female	200.35 (30.9)	1.24 (0.94)	3.27 (1.40)	-	18.3
	Male	119.70 (24.6)	0.96 (0.78)	3.86 (1.87)	-	17.5

Comparison with other studies

Previous authors have compared acoustic analysis obtained by different programmes (Table 3-5) and showed that mean F0 was the measurement least affected by the different algorithms (Amir et al., 2009; Bielamowicz et al., 1996; Karnell et al., 1995). Jitter, shimmer and harmonic-to-noise ratio (HNR) was found to be poorly to moderately reliable when measured between or within programmes (Bough Jr. et al., 1996; Carding et al., 2009; Rabinov et al., 1995). This observed variability is likely to increase in pathological voices (Amir et al., 2009; Bielamowicz et al., 1996; Karnell et al., 1995; Smits et al., 2005). In studies comparing MDVP to other software programmes (Amir et al., 2009; Godino-Llorente et al., 2008), the jitter, shimmer and NHR were highly correlated ($r>7$) but there was a significant difference when pathological voices were measured by different programmes ($p<0.001$) (Amir et al., 2009). However the comparison used Pearson and Spearman correlation tests which are not strictly valid tests to measure realibility of a new instrument (Lee et al., 1989).

Table 3-5: Previous studies inter-software or programme comparison

Study	Test	F0	Jitter %	Shimmer %	NHR or HNR
1995 Karnell et al 5500 vs SEG	Pearson	0.98	0.47	-	-
1995 Karnell et al CSpeech vs SEG	Pearson	1.00	-	0.32	-
1995 Karnell et al CSpeech vs SEG	Pearson	0.98	0.64	-	-
1996 Bielamowicz et al. CSpeech, handmarking CSL and SoundScope	Spearman	0.97 – 0.996	0.33 - 0.61	0.82 – 0.87	HNR 0.23 – 0.86
2005 Smits et al. DrSpeech vs CSL	Pearson	Pearson >0.90	0.69	0.26	HNR 0.74
2008 Godino-Ilorente et al. MDVP vs WPCVox	Pearson	0.98	0.85	0.93	NHR 0.31
2009 Amir et al. MDVP vs Praat	Pearson	>0.90	0.82	>0.70	>0.70

Limitations of the present study

The only acoustic parameter included in the study that had a poor inter-software reliability was NHR. The NHR normative value of the two software programmes was also widely different. This may be due to the different algorithms used. Although MDVP is widely used, there is no consensus among voice experts regarding the 'gold standard' programme for acoustic analysis and hence which algorithm is the 'best'. The intra-software correlation for NHR was good (0.72) but the lower bound 95% CI (0.58), did not reach the accepted level for good reliability. Therefore, clinical application of the NHR measurement derived from OperaVOX must be performed with caution.

Intra-software reliability investigation in this study was performed by recording the voice twice within 15 minutes. The interval was short compared to other studies (Bough Jr. et al., 1996; Carding et al., 2004) possibly contributing to the observed intra-software high ICC.

OperaVOX offers a portable, patient-friendly, cheap and simple means of acoustic analysis, but presently, it does not permit the same depth and breadth of more analysis possible with the static MDVP system. This study shows that neither is 'better' or 'worse' than the other for most measurements in common use, and hence the choice of instrument must depend on the clinical or research application required.

In this paper I removed the most disordered voices from my analysis and attained greater congruity of my results. It may prove that severely disordered voices will ultimately be better evaluated with tools that are adapted for non-linear behavior.

Such tools have begun to be used in conditions such as Cheyne-Stokes respiration and cerebral electrophysiology (Baken, 1996) and are in the very early stages of being applied to pathological voice production (Chunying et al., 2013). That said, at the time of writing of this article, acoustic analysis with more traditional tools is still by and large the technology in use for most laboratory voice analysis.

3.6 Conclusion

Here I showed OperaVOX is statistically comparable to the 'gold standard' (MDVP) for most principal phonatory outcome measures. However, given its portability, low cost and applicability to home or clinic, OperaVOX has greater utility and therefore may be preferred for voice outcome data collection in both settings. Similarly, it may be preferred by patients and clinicians for routine clinical data collection.

Chapter 4 Reliability of video-laryngostroboscopy evaluation in patients with unilateral vocal fold paralysis

4.1 Abstract

Objective: To investigate the inter- and intra-rater reliability of video-laryngostroboscopy evaluation in patients with unilateral vocal fold paralysis (UVFP).

Methods: Videos of video-laryngostroboscopy of 22 patients with UVFP (40 videos) were evaluated by 3 raters (speech and language therapists, SLT) simultaneously but independently. All the videos were duplicated (40 x 2 = 80 videos), anonymised and randomised. Visual perceptual evaluation of the vocal folds' vibratory pattern and structural characteristics (vocal fold position, arytenoid position and vocal fold bowing) was performed. The stroboscopy research instrument (SRI) was used as an evaluation tool to assess the vibratory pattern. The inter- and intra rater reliability was assessed using the kappa-weighted coefficient method.

Results: Three parameters were found to consistently show high kappa (κ) values of substantial to almost perfect inter- and intra-rater reliability: mucosal asymmetry (0.62 to 0.90), right vocal fold periodicity (0.60 to 0.77) and closure duration (0.65 to 0.85). For vocal fold bowing, the reliability was moderate to almost perfect (0.51 to 0.84). Poor to fair reliability ($\kappa < 0.4$) was documented for left amplitude, left vocal fold periodicity, right and left non-vibratory activity, vocal fold position and arytenoid position.

Conclusion: Here, the reliability of video-laryngostroboscopy evaluation in unilateral vocal fold paralysis is reported. Mucosal asymmetry, closure duration and vocal fold bowing appear to be reliable parameters for clinical and research use.

4.2 Introduction

Video laryngostroboscopy is important for the diagnosis of laryngeal disorders and has a role in assessing treatment outcomes (Cutler and Cleveland, 2002; Faure and Muller, 1992). It is widely available and regarded as a vital assessment tool in voice clinics. A survey was done on the use of video-laryngostroboscopy among otolaryngologists of Academy of Otolaryngology-Head and Neck Surgery in which 27.8% of 982 randomly selected members responded (Cohen et al., 2012). The survey found that video-laryngostroboscopy is a common instrument used by otolaryngologists to evaluate the structural abnormality as well as vibration of the vocal folds. However one third of the responders were confused between the purposes of video-laryngostroboscopy and laryngoscopy (Cohen et al., 2012). Laryngoscopy, using rigid or flexible endoscopes is mainly to visualise gross structural characters of the vocal folds such as position, mobility, arytenoid position and the presence or absence of atrophy. video-laryngostroboscopy is an instrument that comprises a laryngoscope (either rigid or flexible) that is connected to a stroboscope and camera to record the examination. Although video-laryngostroboscopy also allows the gross structural evaluation, it has been used to assess precisely the vibratory characteristics of the vocal folds.

Human vocal folds vibrate during phonation that ranged 70 to 500 Hz in males and 130-1,000 Hz in female that is too fast to be seen by naked eyes (Švec and Schutte, 1996). Stroboscopy allows evaluation of dynamic effect of vocal fold vibration-mucosal waves. In Greek, *Strobos* means whirling, while *scopein* means watching or

observing (Faure and Muller, 1992). Video-laryngostroboscopy creates the illusion of slow motion by generating light flashes at the rate that is slightly delayed from the rate of fundamental frequency of the phonation being measured. Thus travelling mucosal waves can be seen showing the vibratory pattern of the vocal folds. It gives an indication of the health of tissue situated between the superficial mucosa (cover) and the vocal muscle (body) (Faure and Muller, 1992; Hirano, 1974). The body and cover are connected by the deep and intermediate layers of the lamina propria. The free motion of the mucosal waves will be affected by any disturbance of the body, cover or intermediate layers. In unilateral vocal fold paralysis (UVFP), flaccid vocal fold muscles reduced the stiffness of the underlying body causing diminution or loss of travelling mucosal waves (Faure and Muller, 1992; Sercarz, 1992). Reinnervation of the muscles that restore the muscle tone or any other surgical interventions that increased the body stiffness may restore travelling waves.

Evaluation of video-laryngostroboscopy findings is typically qualitative: rating by voice experts assessing the visual findings of vocal folds- vocal fold edge, amplitude and symmetry of mucosal wave and glottal closure, as first described by Hirano and Bless (Hirano and Bless, 1993). It is largely dependent on the experience of the rater and that in turn may affect the inter- and intra-rater reliability (Hirano and Bless, 1993). A satisfactory level of inter- and intra-rater reliability in evaluation is important as it reflects the reliability of the tool and for interpretation of the results. Research tools like stroboscopy research instrument (SRI) and Stroboscopy Evaluation Rating Form (SERF) have been developed by Rosen and Poburka respectively to improve the reliability of video-laryngostroboscopy evaluation (Poburka, 1999; Poburka and Bless, 1998; Rosen, 2005).

The SRI is a clinically relevant evaluation tool that was developed by Rosen (Rosen, 2005). Eighteen raters with a mixture of durations of experience demonstrated an intra-rater reliability (Kendall's Tau β) of 0.43 to 0.99 and an inter-rater reliability (intraclass correlation) of 0.11 to 0.65. The reliability was generally higher in the more experienced raters. Since then, a number of trials involving patients with UVFP have used this evaluation tool for video-laryngostroboscopy findings (Morgan et al., 2007; Rosen et al., 2007, 2009; Vinson et al., 2010). However, in most studies that evaluate surgical intervention effect on UVFP, either using SRI or other methods of grading system, the consensus score of two or more raters was sought and the reliability was not explicitly described (Lee et al., 2007; Lorenz et al., 2008; Rihkanen et al., 2004; Wang et al., 2012, 2011).

This study was designed to investigate the inter- and intra-rater reliability of three raters in evaluating video-laryngostroboscopy findings of patients with UVFP, using the SRI and additional measures of structure: vocal fold position, arytenoid position and atrophy (vocal fold bowing).

4.3 Methods

Video selection

Video-laryngostroboscopy was done using flexible nasopharyngolaryngoscope of (VNL-1590STi, KayPentax, USA) attached to a stroboscopy system (RLS 9100B, KayPentax, USA). Forty video-laryngostroboscopy video samples of 22 patients with UVFP (20, left; and 2, right side palsy) presenting to the Royal National Throat Nose Ear Hospital were collected in two ways. First, 15 videos of 15 patients who had been examined from 2008 to 2011 were retrieved from the archives of videos. Second, video samples of 7 patients (25 videos) undergoing laryngeal reinnervation were prospectively collected at different time points, at baseline and follow-ups. Four patients were examined at four time points, and the other three patients at three time points. All of the video samples (15 + 25 = 40) were duplicated twice (40 x 2 = 80) for intra-rater reliability testing. Of 40 videos of the 22 patients, 28 had left, and 12 had right vocal fold paralysis.

The video samples were edited using Windows Movie Maker (Microsoft, USA) for the videos from the archives and iMovie (version 9.0.9, Apple, USA). Videos were edited to a range of 30-120 seconds to avoid raters' fatigue and audio tracks were edited to prevent bias. Hearing the background voice of surgeons on the audio tracks who were doing the video-laryngostroboscopy may provide clues of the identity of patients and time point of examinations whether before or after reinnervation. The video clips were then randomised using Excel (Version 2011, Microsoft, USA) and anonymised. The raters were also unaware whether the videos were recorded before or after laryngeal reinnervation.

Procedures

Evaluation of vocal fold vibration patterns on the video clips was done using the SRI (Rosen, 2005) (Appendix 5). For evaluation of the 'duration of closure', raters were asked to observe the opening and closure of the glottal gap during phonation. It was broken into the categories of predominately closed, half closed/ half opened, predominately opened, or always open. In normal vocal folds, the gap is predominately closed and the opening is brief. Of nine parameters of the SRI, eight items were used in which the 'type of glottal closure' was excluded as to my opinion, it is less of relevance pertaining to UVFP.

The SRI is an assessment tool that mainly used to evaluate dynamic vibration of the vocal folds but not the gross structural characteristics. Therefore, evaluation of the paralysed vocal folds' gross structural characteristics were also included as below:

1. vocal fold position (0, median; 1, paramedian; 2, lateral);
2. arytenoid position at rest (0, normal; 1, partial anteromedial; 2, total anteromedial);
3. arytenoid position during phonation (0, normal; 1, partial anteromedial; 2, total anteromedial); and
4. vocal fold bowing (0, normal; 1, mild bowing; 2, moderate bowing; 3, severe bowing).

These structural characteristics were chosen based on informal discussions with ENT consultants, and the previous studies that document common features of UVFP (Crumley, 1994; Harries and Morrison, 1996; Inagi et al., 1997b; Rosow and Sulica, 2010; Woodson, 1993). These findings frequently contribute to the management of symptomatic patients with UVFP particularly in deciding surgical options.

Three specialist speech and language therapist raters (rater-1, rater-2 and rater-3) evaluated the video clips. They had limited or no prior experience of SRI, but had used video-laryngostroboscopy in voice clinics for over 10 years. A meeting was first arranged to familiarise raters with a few video clips which were then evaluated using the SRI. After the meeting, the 80 pre-prepared edited video clips were evaluated by the raters simultaneously but independently.

Statistical analysis

Inter- and intra-rater reliability for each of the SRI parameter (except type of glottal closure) and gross structural characteristic items were assessed using weighted kappa coefficient. For inter-rater reliability, a pair of raters was assessed at a time. The reliability is considered substantial for kappa values of 0.61 to 0.8 and moderate for kappa values of 0.4 to 0.6 (Landis and Koch, 1977).

4.4 Results

Kappa values for the inter- and intra-rater reliability are summarised in (Table 4-1) and (Table 4-2). Substantial to almost perfect inter-rater reliability for each pair of raters $\kappa = 0.60$ to 0.9 was consistently documented for mucosal wave asymmetry (0.83 to 0.9), right vocal fold aperiodicity (0.61 to 0.77), left vocal fold aperiodicity (0.70 to 0.73), right vocal fold non-vibratory activity (0.69 to 0.85) and duration of closure (0.75 to 0.90). For the rest of the items, at least one of the pair of raters showed moderate inter-rater reliability of $\kappa < 0.6$ except for the left vocal fold non-vibratory activity.

For intra-rater reliability, substantial reliability was demonstrated for mucosal asymmetry (0.62 to 0.66), right vocal fold periodicity (0.60 to 0.65), duration of closure (0.65 to 0.85) and vocal fold bowing (0.64 for rater-2, 0.78 for rater-3). However for rater-1, the inter-rater reliability was moderate (0.51). For the rest of the parameters, κ was < 0.6.

Table 4-1: Summary of kappa values for inter-rater reliability

	Items	R1 vs R2	R1 vs R3	R2 vs R3
SRI	mucosal wave asymmetry	0.90**	0.87**	0.83**
	R Amplitude	0.71*	0.66*	0.59
	L Amplitude	0.60*	0.77*	0.70*
	R periodicity	0.71*	0.77*	0.61*
	L periodicity	0.71*	0.73*	0.70*
	R non-vibratory activity	0.85**	0.69*	0.82**
	L non-vibratory activity	0.00	0.58	0.00
	Closure duration	0.85**	0.90**	0.75*
Structural	Vocal fold position	0.52	0.67*	0.68*
	Vocal fold bowing	0.56	0.68*	0.84**
	Arytenoid position during phonation	0.70*	0.83**	0.53

Table 4-2: Summary of kappa values for intra-rater reliability

	Items	R1	R2	R3
SRI	mucosal wave asymmetry	0.66*	0.62*	0.64*
	R Amplitude	0.59	0.49	0.46
	L Amplitude	0.30	0.37	0.30
	R periodicity	0.61*	0.65*	0.60*
	L periodicity	0.34	0.46	0.28
	R non-vibratory activity	0.30	0.49	0.65*
	L non-vibratory activity	0.13	0.00	0.05
	Closure duration	0.83**	0.65*	0.82**
Structural	Vocal fold position	0.39	0.16	0.55
	Vocal fold bowing	0.51	0.78*	0.64*
	Arytenoid position during phonation	0.35	0.28	0.47

4.5 Discussion

Overview

Video-laryngostroboscopy is useful in clinical practice as well as in research as it may alter choice of therapeutic intervention for voice disorders, as well monitor treatment effects (Cutler and Cleveland, 2002). It enables visualisation of vibration patterns of the vocal folds mucosa that cannot be seen by naked eyes on laryngoscopy. However, video-laryngostroboscopy also allows examination of the structural characteristic of vocal folds that is pertinent in diagnosis of vocal fold paralysis and in assessing the vocal fold position, arytenoid prolapse, vocal fold atrophy (bowing of the vocal fold's edge) and glottal gap. All of these may influence treatment decision (Rosow and Sulica, 2010).

In UVFP, due to discrepancy in the vocal folds' vertical height, tensing capability and muscle tone, and inability to make a firm contact with the opposite normal vocal fold, the mucosal waves are commonly asymmetrical with a significant glottal gap (Faure and Muller, 1992; Sercarz, 1992). The excursion (amplitude) and velocity of mucosal waves on the paralysed side may appear reduced (Sercarz, 1992). Video-laryngostroboscopy findings have been shown to correlate with acoustic and perceptual analysis of voice in which the G factor (overall dysphonia) for the GRBAS scale had the strongest correlation ($r = 0.68$ to 0.88) (Uloza et al., 2005). Video-laryngostroboscopy has been used in monitoring the effect of surgical treatments in patients with UVFP (Lee et al., 2007; Lorenz et al., 2008; Morgan et al., 2007; Paniello et al., 2011; Rihkanen et al., 2004; Rosen et al., 2007, 2009; Wang et al., 2012, 2011). However, the reliability of the evaluation is frequently not reported explicitly in these reports. Hence, the present study was carried out to

investigate the extent of inter- and intra-rater reliability in evaluating the video-laryngostroboscopy findings in patients with UVFP.

Strength of the study

This is the first study that rigorously investigates the inter- and intra-rater reliability in evaluating the vibratory pattern and vocal fold structural characteristics of UVFP. The results may reflect the reliability of parameters and suggests which parameters might be most reliable in clinical trials. SRI was chosen instead of SERF because the SRI is more clinically relevant and less complicated. This is important to avoid raters fatigue that may affect the result of the present study. The SRI also has been used in the previous studies of surgical effectiveness in UVFP.

Summary of key findings

The present study looked at the inter- and intra-rater reliability in evaluating video-laryngostroboscopy findings using the SRI found that two parameters had consistently substantial to almost perfect reliability: mucosal asymmetry and duration of closure. For the vocal fold paralysis structural evaluation, only vocal fold bowing had consistent substantial inter- and intra-rater reliability although for one of the raters the intra-rater reliability was only moderate.

Comparison with previous work

A number of reliability studies have been carried out to investigate the reliability of UVFP evaluation tools. In these, inter-rater reliability ranged from low to good and some of the items measured vary between sides. There was no consistency in the types of vocal fold disorder examined and recorded using video-laryngostroboscopy.

Rosen found that, of nine parameters of the video-laryngostroboscopy evaluation tool (SRI), four depicted an intraclass correlation coefficient (ICC) of acceptable to substantial reliability: left vocal fold amplitude (0.61), left non-vibratory portion (0.68), closure pattern (0.59), and closure duration (0.58) (Rosen, 2005). Although the left vocal fold amplitude and left non-vibratory pattern was acceptable but the right side of each was 0.25 and 0.11 respectively. Nawka and Konerding investigated the reliability of another video-laryngostroboscopy evaluation tool developed by Poburka and reported that the ICC ranged from 0.32 to 0.71. Parameters that depicted an ICC of acceptable to subtantial reliability were left mucosal wave (0.59), right non-vibrating portion (0.57), and left and right vocal fold straightness (0.71 and 0.64) (Nawka and Konerding, 2012). The contralateral parameters of the mucosal wave and non-vibrating portion did not reach the acceptable reliability. Subsequently, Yiu et al. re-examined the reliability of amplitude, supraglottic activity and glottal closure and added another parameter, lesion size rating. The study reported that the inter- and intra-rater reliability for lesion size was good (ICC, 0.75 to 0.81), antero-posterior supraglottic activity was substantial (ICC, 0.64) and glottal closure was substantial (κ, 0.64) (Yiu et al., 2013). Uloza et al. evaluated inter- and intra- rater reliability of video-laryngostroboscopy parameters that were measured using a 100-mm long visual analogue score and documented that the reliability of glottal closure, regularity, and symmetry of vibration was substantial to almost perfect (Uloza et al., 2013).

Rosow et al. evaluated the inter- and intra-rater reliability of 12 parameters of laryngoscopy appearance of structural characteristics of 22 patients with vocal fold

paralysis. The study reported that glottic insufficiency, vocal fold bowing and salivary pooling depicted the highest inter-rater reliability κ: 0.55, 0.49 and 0.45 respectively although they are of moderate reliability (Rosow and Sulica, 2010).

The present study was the first to rigorously evaluate the reliability of video-laryngostroboscopy parameters (SRI) and structural characteristics of vocal folds in patients with UVFP. These parameters were chosen because they are pertinent in vocal fold paralysis assessment and frequently measured in published studies investigating the efficacy of surgical interventions. Results of the present study are similar to Rosen's for closure duration, to Uloza's for symmetry of vibration and to Rosow's for vocal fold bowing.

Limitations of the present study

There are several factors that may confound visual perceptual evaluation: laryngoscopy procedures, parameters for evaluation, evaluation tool used, experience of the raters and amount of training applied(Hirano and Bless, 1993; Yiu et al., 2013).

The raters involved in the present study, although very experienced in assessing video-laryngostroboscopy, had limited experience in using SRI. There was no formal training given prior to the familiarisation meeting and this may skew results. Therefore, it is necessary to provide formal training to raters as well as to encourage them to use the tool in their daily clinical practice.

Evaluation of dynamic vibration of vocal folds is relatively difficult. The evaluation may be inaccurate if the tip of the laryngoscope is angled hence amplitude may seem larger on one side than the other. This angulation also may affect the laryngoscopy appearance of vocal fold position and arytenoid position. A careful manoeuvring of the laryngoscope to avoid this situation may improve the reliability of video-laryngostroboscopy evaluation.

Kappa weighted coefficient is a standard test of inter- and intra-rater reliability or ordinal data (Cohen, 1968; Kottner et al., 2011; Viera and Garrett, 2005; Yiu et al., 2013). However the results may be dependent on the prevalence of pathology or abnormality under consideration. For small prevalence findings, a low kappa value may not necessarily reflect the overall agreement of the raters (Viera and Garrett, 2005). In the present study, 20 videos belonged to 7 patients that had been taken at different time points post-laryngeal reinnervation. Therefore, normal or mildly abnormal vibration patterns or vocal fold structural findings as direct results of the intervention may introduce bias and affect kappa values. Future studies should include patients with a wider variety of UVFP severity.

Clinical impact

The present study supports the use of video-laryngostroboscopy in evaluating the mucosal wave asymmetry, duration of closure and vocal fold bowing in patients with UVFP. These parameters are important in indicating return of voice function by closing the glottal gap as well as the preservation of pliability of the vocal fold mucosa, for example that following laryngeal reinnervation.

4.6 Conclusion

Here, I demonstrated the reliability of video-laryngostroboscopy as a tool for evaluating UVFP. A substantial inter- and intra-rater reliability was shown for the mucosal wave asymmetry, duration of closure and vocal fold bowing of the video-laryngostroboscopy parameters that appear to be reliable parameters for clinical and research use.

Chapter 5 Denervation changes on T2-MRI of vocal fold paralysis: a repeatability study

5.1 Abstract

Objective: This study investigated the repeatability and reproducibility of T2-weighted MRI (T2-MRI) to depict laryngeal muscles' – thyroarytenoid (TA) and posterior cricoarytenoid (PCA), muscle bulk and signal changes in patients with UVFP.

Methods: T2-MRI was performed on five healthy volunteers and nine UVFP patients. Images were acquired in axial and coronal planes to depict TA and PCA (resolution, 300x297; slice thickness, 2.5mm). Five participants from each group had a second scan. All patients underwent laryngeal electromyography (LEMG) in which recruitment of motor units were graded by a neurophysiologist, unaware of MRI findings. Region of interest (ROI) was placed independently by two readers on:

1) TA coronal to measure ***TA bulk*** cm^2
2) PCA axial to measure ***PCA bulk*** cm^2
3) TA axial to measure ***TA signal-intensity***
4) PCA axial to measure ***PCA signal-intensity***

Within subject repeatability and inter- and intra-reader reproducibility was assessed using intraclass correlation test (ICC). Comparison of metrics between groups was done using Mann Whitney test.

Results: TA signal intensity were significantly higher, and TA bulk and PCA bulk were significantly smaller in paralysed muscles than normal muscles (p=0.03 to 0.01). TA signal intensity and PCA bulk were significantly correlated with the motor

unit recruitment on LEMG. Adequate repeatability and reproducibility (ICC=0.63 to 0.83) was persistently demonstrated for TA bulk and signal-intensity.

Conclusion: I demonstrated the repeatability and reproducibility of T2-MRI in depicting denervation changes in UVFP. Signal changes on TA muscles correlate with electrophysiological results on LEMG and may be useful in monitoring treatment outcome of laryngeal reinnervation in UVFP patients.

5.2 Introduction

The larynx is the main route for human social interaction and critical for breathing, swallowing and airway protection. Loss of laryngeal function therefore has major implications. Muscles within the vocal folds of the larynx are innervated mainly by the recurrent laryngeal nerve (RLN), a branch of the vagus nerve. When the nerve supply to one vocal fold is interrupted, unilateral vocal fold paralysis (UVFP) ensues. Patients with UVFP typically experience hoarse, weak and an easily tired voice. Quality of life is significantly reduced and social and employment opportunities are compromised (Fang et al., 2008; Smith et al., 1995). Aspiration of saliva and food, and airway obstruction are sometimes problems also.

The aim of treatment for UVFP is to achieve optimum vocal fold closure for voice production. One of surgical interventions that have been proposed is laryngeal reinnervation. It restores innervation to the larynx by 'borrowing' the activity of other motor nerves, commonly the ansa cervicalis. As such it re-establishes the tone and bulk of the vocal folds for speech and re-enables pitch control resulting in normal or near normal voice (Wang et al., 2011).

Evaluation of laryngeal muscle characteristics before and after reinnervation is important to confirm eligibility as well as monitor treatment effect. Reinnervation surgery should not be performed without confirming the status of muscle paralysis. Sometimes mechanical fixation of the cricoarytenoid joint may be mistakenly diagnosed as vocal fold paralysis and this need to be ruled out (Connor et al., 2006; Romo and Curtin, 1999). The present investigation for this purpose is laryngeal electromyography (LEMG).

LEMG is commonly used to assess the neurological status of paralysed laryngeal muscles. In denervated laryngeal muscles, the electromyography shows absence or reduction of normal motor unit action potentials (MUAP) and recruitment during voluntary contraction with presence or absence of spontaneous activity. It is useful in showing diagnosis and prognosis of vocal fold paralysis although it is based on class IV evidence. It has been assumed as the gold standard as it is the only tool used to provide physiological evidence to reinnervation or denervation of laryngeal muscles (Blitzer et al., 2009). However, it requires considerable operator skill, is invasive and muscles are not directly visualized during needle positioning. It requires an experienced otolaryngologist, electromyographer and a neurophysiologist to get a reliable result (Volk et al., 2012). The otolaryngologist has to be someone who familiar with laryngeal anatomy and has been doing botulinum injection to the larynx by external approach regularly. Availability of LEMG is also limited; in a survey done among ENT consultants in the UK, only 20% has the access to LEMG whereas 100% had access to MRI (unpublished data).

MRI has been shown useful in depicting denervation changes – increased signal intensity and reduced bulk, in muscles. Signal changes were shown on the T2-weighted MRI (T2-MRI) during acute, sub-acute and chronic phase of denervation (Kamath et al., 2007). Therefore, it may provide a non-invasive alternative or an adjunct to LEMG to directly assess laryngeal muscle bulk and signal intensity changes associated with paralysis (Connor et al., 2006; Romo and Curtin, 1999). Early anatomical 1.5T MRI has been used to assess laryngeal muscle bulk and fatty change in patients with vocal fold paralysis; but no quantitative metrics were

evaluated and no correlation with a corroborating standard of reference was performed (Romo and Curtin, 1999).

This study investigated the reliability and validity of T2-MRI to depict laryngeal muscles' – thyroarytenoid (TA) and posterior cricoarytenoid (PCA), muscle bulk and signal intensity in patients with UVFP.

5.3 Methods

Local Ethics Committee approval was obtained (reference number: 11/LO/0583) for the study.

Study overview

This study investigates the reliability and validity of the T2-MRI for quantitatively evaluating muscle bulk and signal changes of TA and PCA in UVFP alongside healthy controls. Four outcomes were of interest in this study:

1. Within subject repeatability of T2-MRI metrics measurement between first and second scan at two separate occasions (inter-scan).
2. Inter- and intra-reader reproducibility of T2-MRI metrics measurements.
3. Comparison of T2-MRI metrics between healthy volunteers and UVFP patients.
4. Correlation of T2-MRI metrics with LEMG scores.

Participant selection

Healthy controls: Adult volunteers without any current or previous history of voice problems (excluding previous upper respiratory tract infection). Five healthy volunteers (2 female, 3 male) with age ranged from 28 to 41 old were included in this study.

Unilateral vocal fold paralysis patients: 9 UVFP patients (7 female, 2 male) with age ranged from 19 to 62 years old were recruited. UVFP was identified by fibre-optic nasopharyngolaryngoscope of (VNL-1590STi, KayPentax, USA) and confirmed by LEMG which was done after the scan. The causes were secondary to

iatrogenic or idiopathic causes. All patients had left vocal fold paralysis for more than 6 months and had never undergone any medialisation procedure.

Five participants from each of the volunteer and patient groups had second visit scan within four weeks to assess the inter-scan repeatability.

MRI protocol

Imaging was performed on a 3T Philips Achieva TX MRI scanner (Philips Healthcare, The Netherlands). Sense Flex S coils of 11cm diameter were placed at the side of the neck covering the anterior part of the larynx below the angle of the mandible, while the participant lied supine on the scanner bed with the head resting on a headrest. The coils were held in place by putting a neck strap and sandbags at the side of the head and neck to minimize movement. The participants were asked not to move their neck or jaw, clear their throat or swallow during the image acquisition period. The T2-MRI protocol is shown in Table 5-1.

Table 5-1: This table shows T2-MRI protocol for axial and coronal view

T2-MRI protocol	Coronal	Axial
Sequence Name	SE-TSE	SE-TSE
Repetition time (ms)	6789	3903
Echo time (ms)	120	100
Image resolution	0.5x.5x	0.5x0.5
Slice thickness	2.5	2.5
Number of slices	14	18
Flip angle	90	90
Field of view (FOV) mm	140x140	140x140
Total acquisition time	5 min 32 secs	7 min 32 secs

LEMG protocol

UVFP patients had LEMG performed within 4 weeks after the MRI. While patients lying on the couch with neck slightly extended, a laryngologist (MAB) inserted a monopolar needle into the TA muscles through the cricothyroid membrane while an electromyographer operated the electromyography machine. The PCA muscles were not studied due to technical challenges in accessing the muscle caused by deep location of the muscle. The LEMG findings were captured and saved. The LEMG tracing recording was then converted to an executable file (exe.file) that later rated by a consultant neurophysiologist who was unaware of the T2-MRI findings. The LEMG findings were evaluated according to Koufman classification whereby the item of interest in the present study is the recruitment of motor units (Koufman et al., 2001). The motor unit recruitment was graded from 0 (no recruitment) to 4 (normal recruitment). LEMG results of the TA were inferred for PCA as both are innervated by the RLN

Image analysis

Images were downloaded to OsiriX (open-source software, Geneva, Switzerland) (Rosset et al., 2004) workstation for analysis. They were then quantitatively analysed.

Two readers independently made the T2-MRI metrics measurements to assess inter-reader reproducibility. The first reader, an Otolaryngology surgeon who has been trained in laryngology for 4 years and the second reader, a consultant radiologist experienced in body MRI (10 years). The first reader made a second measurement after a month from the first measurement to assess intra-reader reproducibility. The

muscle bulk of left and right, TA and PCA were evaluated by measuring the muscle area (cm^2) on the slice that depicts the biggest muscle bulk by drawing a region of interest (ROI) on the largest cross-section of the muscle body. The TA bulk was measured on the coronal plane images (Figure 5-1) whereas for the PCA muscle measurements were made on axial images (Figure 5-2). Mean signal intensity of the muscles was obtained from the ROI drawn on the axial plane for both muscles (Figure 5-3). The signal intensity was then normalised against individual patient skeletal neck muscle by calculating the ratio of the laryngeal muscles to the left sternocleidomastoid muscle signal intensity.

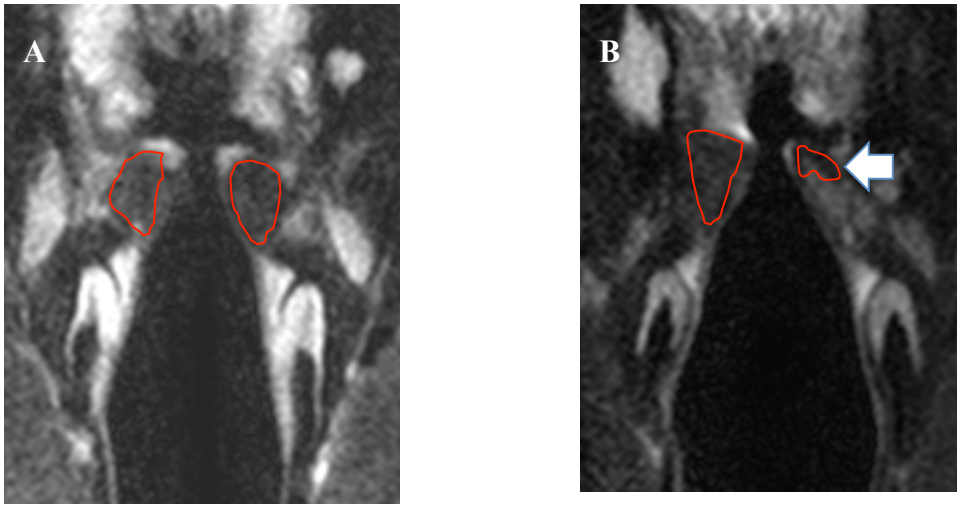


Figure 5-1: This figure shows T2-MRI of larynx images in coronal view showing the main cross-sectional bulk of thyroarytenoid muscles of a: A) healthy volunteer; and B) unilateral vocal fold paralysis patient. The arrow shows left side paralysis in which the bulk is smaller than the normal side.

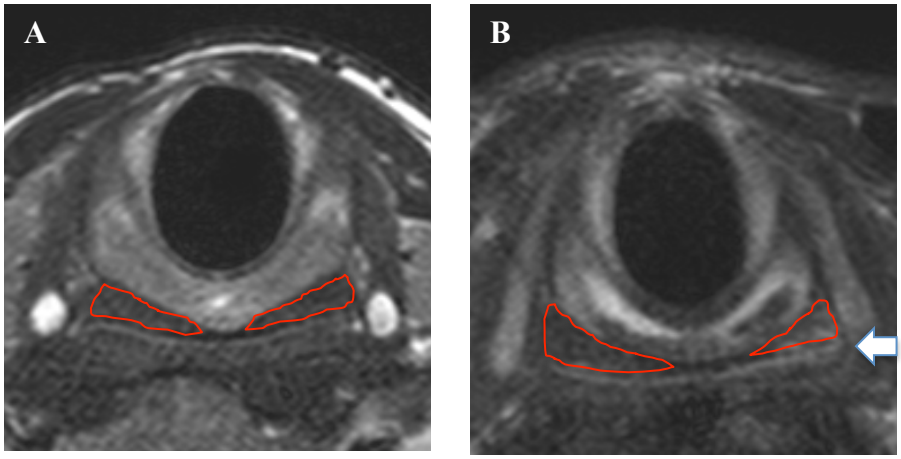


Figure 5-2: This figure shows T2-MRI of larynx images in axial view showing posterior cricoarytenoid muscles of a: A) healthy volunteer; and B) unilateral vocal fold paralysis patient. The arrow shows left side paralysis in which the bulk is smaller and the signal intensity is higher than the normal side. The region of interests are drawn in redlines.

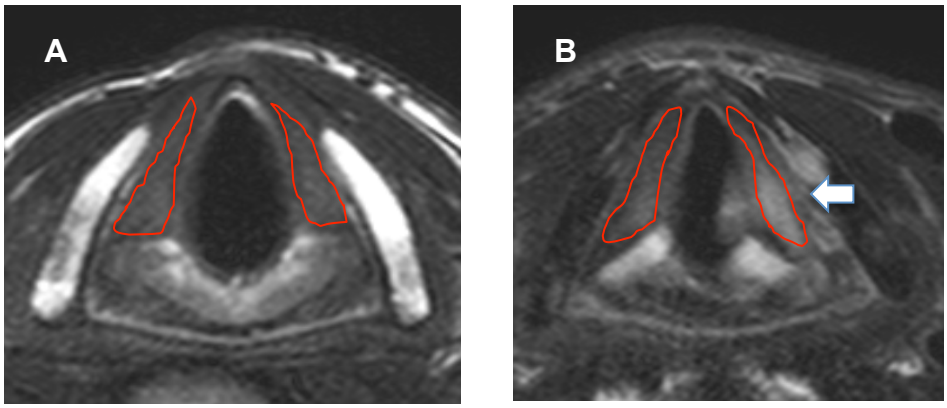


Figure 5-3: This figure shows T2-MRI of larynx images in axial view showing thyroarytenoid muscles of a: A) healthy volunteer; and B) unilateral vocal fold paralysis patient. The arrow shows left side paralysis in which the signal intensity is higher than the normal side. The region of interests are drawn in redlines.

Statistical analysis

Data of this study is presented in median and interquartile (IQR) range.

1. Within subject repeatability of T2-MRI metrics measurement (inter-scan) was assessed using intraclass correlation coefficient test (ICC)- two-way mixed, consistency.
2. Inter- and intra-reader reproducibility of T2-MRI metrics measurements was also assessed using the ICC- two-way mixed, absolute agreement.
3. T2-MRI metrics were compared:
    a. within volunteers, left and right TA and PCA;
    b. within patients, paralysed and normal TA and PCA;
    c. between volunteers and patients, normal TA and PCA; and
    d. between volunteers normal and patients, paralysed TA and PCA.
4. Correlation of T2-MRI metrics with LEMG scores was assessed using Spearman rank correlation.

Comparison between muscles within subjects was performed using the Wilcoxon signed rank test. The Mann Whitney test was used for comparison across volunteer and patient groups. ICC values ranged from 0 (not repeatable or reproducible) to 1 (perfect repeatability or reproducibility). The ICC has to be at least 0.6 for the measurements to be clinically useful (Chinn, 1991). Repeatability and reproducibility of the measurements was considered good for ICC of similar to 0.75 and moderate for ICC of between 0.5 and 0.75 (Portney and Watkins, 1993).

5.4 Results

The median and interquartile ranges (IQR) for the left and right TA bulk, PCA bulk, TA- and PCA signal intensity of healthy volunteers and patients are presented in Table 5-2.

Table 5-2: This table shows the summary of anatomical parameters between healthy volunteers and UVFP patients. Where medians are shown, brackets represent interquartile ranges.

	Volunteer	Patient
TA bulk		
Left	0.30 (0.22, 0.32)	0.14 (0.12, 0.26)
Right	0.29 (0.22, 0.37)	0.32 (0.23, 0.38)
PCA bulk		
Left	0.33 (0.28, 0.42)	0.18 (0.16, 0.26)
Right	0.33 (0.25, 0.42)	0.36 (0.31, 0.46)
TA signal intensity		
Left	1.98 (1.71, 2.31)	2.84 (2.17, 3.61)
Right	1.99 (1.84, 2.39)	2.06 (1.58, 2.44)
PCA signal intensity		
Left	1.49 (1.22, 1.80)	1.69 (1.60, 2.41)
Right	1.62 (1.32, 2.03)	1.51 (1.39, 1.93)

Within subject repeatability

PCA bulk and TA signal intensity T2-MRI metrics demonstrated good within subject repeatability with ICC values of 0.79 and 0.83 respectively. TA bulk and PCA signal intensity had moderate repeatability (0.52 and 0.63, respectively). However the PCA signal intensity was still clinically useful.

Inter- and intra-reader reproducibility

TA bulk, TA signal intensity and PCA signal intensity T2-MRI metrics showed moderate inter-reader reproducibility with ICC values of 0.65, 0.71 and 0.73 respectively, adequate to be clinically useful. PCA bulk had moderate reproducibility but not clinically useful (0.57).

For intra-reader reproducibility, the TA signal intensity had good repeatability (ICC, 0.76) and the TA bulk and PCA signal intensity had moderate repeatability with ICC values of 0.71and 0.74 respectively, adequate to be clinically useful (Table 5-3). The PCA bulk had moderate intra-reader reproducibility but not clinically useful (0.46).

Table 5-3: This table shows within subject repeatability, and inter- and intra-reader repeatability results.

T2-MRI metrics	Within subject	Inter-reader	Intra-reader
TA bulk	0.63 (0.27, 0.83)	0.65 (0.20, 0.84)	0.71 (0.01, 0.90)
PCA bulk	0.79 (0.54, 0.91)	0.57 (0.11, 0.79)	0.46 (-0.10, 0.77)
TA signal intensity	0.83 (0.62, 0.93)	0.71 (0.45, 0.85)	0.76 (0.60, 0.86)
PCA signal intensity	0.40 (-0.04, 0.71)	0. 73 (0.56, 0.84)	0.74 (0.57, 0.85)

ICC values are presented in the table

Comparison between healthy volunteers and UVFP patients

Within volunteers, there was no significant difference for TA bulk, PCA bulk and TA signal intensity T2-MRI metrics between left and right muscles (p values ranged from 0.27 to 0.50) except for PCA signal intensity ($p = 0.04$).

Within patients, the TA bulk and PCA bulk were significantly smaller ($p=0.01$ and $p=0.01$) and the TA signal intensity was significantly higher on the paralysed side than the normal side. For PCA signal intensity, there was no significant difference between the paralysed and the normal side ($p = 0.21$) (Figure 5-3)

For between volunteers and patients' normal muscles, there was no significant difference for all the T2-MRI metrics ($p=0.46$ to 0.96).

For between volunteers' normal and patients' paralysed muscles, the TA bulk and PCA bulk were significantly smaller ($p=0.03$ and $p=0.01$) and the TA signal intensity was significantly higher, on the paralysed side than the normal side. For PCA signal intensity, there was no significant difference between the paralysed and the normal side ($p=0.32$).

Table 5-4: This table shows results of comparison between healthy volunteers and UVFP patients.

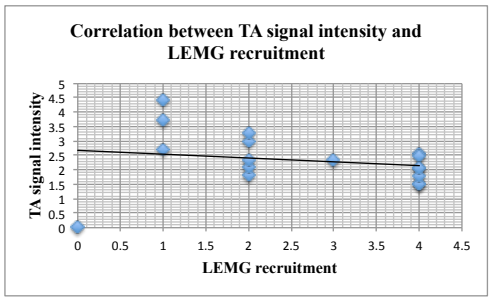
	Within patient	Within volunteer	Volunteers vs patients, (normal)	Volunteers (normal) vs patients (paralysed)
TA bulk	0.01	0.28	0.46	0.03
PCA bulk	0.01	0.27	0.46	0.01
TA signal intensity	0.02	0.50	0.96	0.012
PCA signal intensity	0.21	0.04	0.57	0.32

p value of the Wilcoxon sign rank test is presented in the table

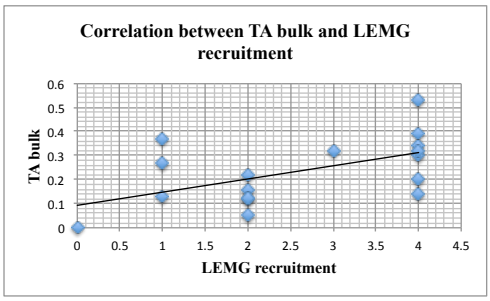
Correlation between T2-MRI metrics and LEMG recruitment grade

There was strong negative correlation between LEMG recruitment motor unit and TA signal intensity ($r = -0.70$, $p = 0.003$) that indicates a reduction in motor unit recruitment is related to an increase in TA signal intensity. TA bulk had moderate positive correlation ($r = 0.40$, $p = 0.12$) with motor unit recruitment.

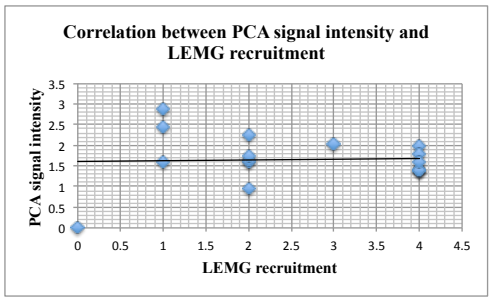
Strong positive correlation was demonstrated between LEMG recruitment grade and PCA bulk ($r = 0.66$, $p = 0.006$) that shows a reduction in motor unit recruitment is related to a reduction in PCA bulk. PCA signal intensity had moderate negative correlation ($r = -0.42$, $p=0.11$) with motor unit recruitment.



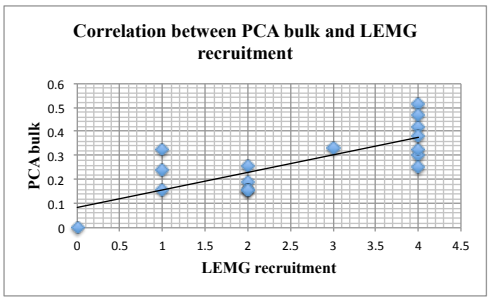
r = -0.70, p = 0.003



r = 0.40, p = 0.12



r = -0.42, p=0.11



r = 0.66, p = 0.006

Figure 5-4: This figure shows graphs of correlation between T2-MRI metrics and laryngeal electromyography recruitment. r: correlation value, p: p value

5.5 Discussion

Overview

LEMG is the present available gold standard for evaluation of denervated laryngeal muscles and monitoring of reinnervation. Evaluation of laryngeal muscles neurological status is important to decide timing and type of surgical intervention (Koufman et al., 2001). It is particularly important before a laryngeal reinnervation procedure to confirm the muscles' neurological status and after the operation to measure treatment effect supporting the voice outcome measures. However, LEMG is invasive, operator dependent and not routinely practiced by laryngologists (Volk et al., 2012). The LEMG evaluation is widely qualitative that is subject to expertise and familiarity although there were studies been done to investigate the utility of objective evaluation (Smith et al., 2012; Statham et al., 2010). MRI may be a useful non-invasive method in objective assessment of neurological status providing evidence of denervation and reinnervation of laryngeal muscles. It has been shown that changes on MRI mirrored electrophysiological changes on electromyography (Küllmer et al., 1998).

MRI has been useful in demonstrating denervation changes of skeletal muscles affecting the morphology and physiology (Kamath et al., 2007; Koltzenburg and Yousry, 2007) and predicting prognosis (Yamabe et al., 2008). It can objectively depicts evidence of denervation and reinnervation by measuring the muscle volume and signal changes on T2-MRI (Bendszus et al., 2002; Kikuchi et al., 2003; Viddeleer et al., 2012; Yamabe et al., 2008). In peripheral nerve neuropathies, the hyperintense signal on T2-MRI is due to an increase in extracellular fluid secondary to oedema or increased in vascularity and fatty degeneration (Kikuchi et al., 2003;

Polak et al., 1988; Wessig et al., 2004) depending on the duration of nerve injury. The abnormal signal changes on T2-MRI correlated with prognosis in rats (Yamabe et al., 2008) and with hand function in human (Viddeleer et al., 2012).

Synopsis of key findings

The findings in the present study demonstrated that the TA signal intensity was significantly higher, and the TA bulk and PCA bulk were significantly smaller in the paralysed than the normal muscles. The TA signal intensity and TA bulk also had strong significant correlation and moderate correlation respectively with the motor unit recruitment. For PCA bulk and PCA signal intensity, strong significant correlation and moderate correlation respectively with the motor unit recruitment were shown. These results indicate that the signal intensity increased as recruitment of motor units reduced and a reduction of muscle bulk on T2-MRI was related to reduction of motor unit recruitment. I expect these results are likely to demonstrate further statistical significance with higher number of patients included in the study.

The present study also demonstrates adequate within subject repeatability and inter- and intra-reader reproducibility of the T2-MRI metrics measurement except for the PCA bulk. Among the T2-MRI metrics, TA bulk and TA signal intensity consistently depicts adequate repeatability and reproducibility by which the TA signal intensity had good within subject repeatability and intra-reader reproducibility.

Comparison with previous work

MRI in laryngeal paralysis is widely used to investigate its etiologies (Vachha et al., 2013). Sakai et al. investigated normal laryngeal structures in 62 participants using

1.5T machine and had shown that MRI is efficient in demonstrating laryngeal muscles (Sakai et al., 1990). MRI able to clearly depict denervation changes on the TA and PCA muscles (Borges, 2010; Connor et al., 2006). A retrospective study was done to evaluate the ability of MRI in depicting denervation atrophy of laryngeal of laryngeal muscles (Romo and Curtin, 1999). This study involved 20 UVFP patients in whom four of them had both CT scan and MRI. Denervation atrophy was detected in high frequency 70 – 90%. They concluded that PCA muscle atrophy may be a useful indicator of RLN paralysis when other findings are subtle or equivocal.

However studies to correlate MRI changes with electrophysiological changes in denervated laryngeal muscles is lacking in the literature. To my knowledge the present study is the first to quantify the MRI changes on T2-MRI objectively and rigorously test the repeatability and validity against the electrophysiological changes on LEMG.

Limitation of the present study

T2-MRI images are susceptible to motion. Participants were not allowed to clear their throat or swallow during acquisition period to ensure good quality of images. Some UVFP patients may find it difficult especially when they have pooling of saliva due to swallowing issues. One dataset was excluded due to the motion artifact. This issue may be able to overcome by improving the time acquisition in the protocol.

The TA muscles are commonly injected with substances for medialisation procedure that may alter the signal intensity. Therefore the use of T2-MRI in signal intensity

measurement of the TA may be limited in those patients who had received such procedure by which the high signal intensity can be attributed to the injection instead of fatty change or oedema. In this situation, PCA muscles may be used as an indicator (Romo and Curtin, 1999). However, PCA bulk had the lowest reliability and PCA signal intensity was not significantly different between the paralysed and normal muscles.

PCA muscles were not studied for the LEMG due to technical difficulties in locating the muscle. The TA muscles' LEMG results were used instead as both muscles were supplied by the same nerve. In a study that investigate spontaneous reinnervation in cats, reinnervation was documented to be more robust in TA muscles – adductor, than the PCA muscle – abductor (Woodson, 2007). Therefore correlation between PCA muscles metrics and the LEMG results need to be interpreted with caution.

T2-MRI may not be suitable in in certain group of vocal fold paralysis patients such as those who had any screws or implants in the cervical spine or had stainless steel heart valve replacement or pacemaker that are not MRI compatible. If they are compatible, foreign bodies in the cervical spine will cause artifact that will affect quality of images.

This study included small number of patients with limited range of neurological status on LEMG results. Therefore it cannot conclude whether T2-MRI anatomical metrics able to demonstrate the severity of paralysed muscles' neurological status. Future studies should recruit more number of patients with wider temporal range, from acute phase to chronic phase of denervation.

Clinical impact

T2-MRI is useful to depict denervation and monitor effect of reinnervation procedure in UVFP. A successful reinnervation was shown to reverse the high signal changes as shown in other skeletal muscles in animal models (Yamabe et al., 2008). The present study demonstrated TA signal intensity as the most useful parameter that had strong correlation with the motor units recruitment, good within subject repeatability and intra-reader reproducibility, and clinical useful intra-reader reproducibility. Other metrics may be useful as well and a better reliability results may be demonstrated if a bigger number of participants is included. A larger scale study that include more numbers of volunteers and patients is necessary to decide on a normal cut off range.

The ability of the T2-MRI to differentiate between normal and paralysed muscles will help surgeons to decide on patients' eligibility of reinnervation. It confirms paralysis by demonstrating smaller bulk (atrophy) and higher signal intensity on the paralysed side than the opposite normal. It also may exclude total denervation that is an exclusion criterion for reinnervation procedure. The T2-MRI also, potentially able to show increased muscle bulk following reinnervation, which contributes to voice improvement.

5.6 Conclusion

I demonstrated the repeatability and reproducibility of T2-MRI in depicting denervation changes in vocal fold paralysis muscles. Signal changes on TA muscles

had shown to be correlated with electrophysiological results on LEMG that is useful in monitoring treatment outcome of laryngeal reinnervation in UVFP patients.

Chapter 6 Muscle perfusion in laryngeal muscles paralysis: Findings on dynamic contrast enhanced MRI and diffusion weighted MRI

6.1 Abstract

Objective: To investigate the perfusion of thyroarytenoid (TA) and posterior cricoarytenoid (PCA) muscle in healthy volunteers and unilateral vocal fold paralysis (UVFP) patients using dynamic contrast enhanced-MRI (DCE-MRI) and diffusion weighted-MRI (DWI).

Methods: DWI and DCE-MRI were performed on five healthy volunteers and nine UVFP patients. Axial images were acquired to depict TA and PCA. Regions of interest were placed within left and right of TA and PCA of acquired images. For DCE-MRI, time intensity curve (TIC) were generated by plotting signal intensities during and following intravenous contrast administration. TIC slope of initial enhancement (SoE), maximum enhancement (ME) and area under the curve (AUC) at 1 minute (AUC-1) was derived. For DWI**,** apparent diffusion coefficient values **(**ADC) were calculated from low b-values 0, 50, 100. Left and right TA, and PCA MRI metrics within subjects were compared using the Wilcoxon sign rank test. The Mann Whitney test was used for comparison between volunteers and patients.

Results: There was no significant difference of SoE, ME, AUC-1 and ADC, between the left and right of- TA and PCA in volunteers, and between volunteers and patients. However PCA DCE_MRI metrics were significantly higher in patients normal than volunteers. For patients, the paralysed TA and PCA ADC were significantly lower than normal muscle within subjects ($p=0.01$, $p=0.01$) and between patients and volunteer groups ($p = 0.08$, $p = 0.03$). For DCE-MRI, the SoE, ME and AUC-1 ($p=0.01$) was lower in the paralysed PCA within patients ($p=0.01$) and between patient and volunteer groups ($p=0.01$ to 0.03).

Conclusion: This study demonstrates the feasibility of assessing tissue perfusion in laryngeal muscles using DCE-MRI and DWI. The findings confirm that paralysed

TA and PCA muscles have reduced perfusion and that apparently normal contralateral musculature in patients with UVFP has elevated perfusion.

6.2 Introduction

Normal skeletal muscles contain a capillary network that allows replacement of glucose, fatty acid and oxygen consumed in muscular activity as well as removal of lactic acid and carbon dioxide. Peripheral neuropathy causes these capillaries to involute, possibly due to an adaptive mechanism to reduce load in a functionless region or in response to an adverse local niche (Carpenter and Karpati, 1982). There have been reports that this process is followed by a relative increase in perfusion as muscles atrophy (Carpenter and Karpati, 1982; Murakami et al., 1998; Russo et al., 1997) although others suggest a reduced microcirculation demonstrated by progressive reduction of the capillary-to-fibre ratio and capillary density per cross sectional area following denervation (Borisov et al., 2000)(Lu et al., 1997).

Such studies have so far depended on animal models (Kato et al., 2002; Miyamaru et al., 2008; Shindo et al., 1992; Zealear et al., 1994). Evaluation of denervation changes in human thyroarytenoid (TA) and posterior cricoarytenoid (PCA) muscles by histology can only realistically be achieved post-mortem and so changes in perfusion have only been inferred or extrapolated. However, progressive innovation in MRI technology, and in particular diffusion weighted imaging (DWI) and dynamic contrast enhanced magnetic resonance imaging (DCE-MRI), may now present an alternative for evaluating the pathologic properties and physiologic changes within denervated muscles.

DCE-MRI is based on acquisition of serial T1-weighted images prior to, during and following administration of intravenous gadolinium contrast (Lavini et al., 2013; Verstraete et al., 1996; Weber et al., 2007). Changes in the signal intensity of tissues

with time can provide functional information related to capillary density (perfusion), structure and permeability of microvasculature (Lavini et al., 2013; Verstraete et al., 1996). DCE-MRI has been used to investigate tissue microcirculation differences between malignant and benign tumours, post-radiation changes and evaluate the effects of therapeutic interventions (Knopp et al., 2001; Punwani, 2011).

Another MRI method for measuring tissue perfusion is DWI. This relies on the sensitivity of MRI to water mobility. Within cellular and interstitial spaces DWI provides a measure of 'Brownian motion' of tissue water (Le Bihan et al., 1991); but it can also be made sensitive to intra-vascular 'bulk motion' of water (Padhani et al., 2009). By acquiring diffusion images using low b-values the diffusion signal is sensitised to changes in tissue perfusion (Yanagisawa et al., 2009a, 2009b). 'b-value' is a diffusion gradient strength applied to T2-weighted spin-echo sequences that causes dephasing and rephasing of water molecules. Signal loss during rephasing is achieved due to water protons move out of plane. Diffusion gradient strength of 0 to 100s/mm^2 is considered low b-values and of more than 100s/mm^2 is considered high b-values. However, studies of denervated skeletal muscle perfusion using the low b-values DWI are lacking.

To my knowledge, there have been no published studies on perfusion measurement within denervated laryngeal muscles using DWI and DCE-MRI. I aimed to investigate the perfusion of TA and PCA in healthy volunteers and unilateral vocal fold paralysis patients using DWI and DCE-MRI.

6.3 Methods

Ethics statement.

The local ethics committee approved this study and informed consent was obtained from all participants.

Participant selection

Healthy controls: Five healthy volunteers (2 female, 3 male) without any current or previous history of voice problems (excluding previous upper respiratory tract infection) aged between 28 and 41 years' old were recruited.

Unilateral vocal-fold paralysis patients: Nine UVFP patients (7 female, 2 male) aged between 19 and 62 years were recruited. UVFP was diagnosed by fibre-optic nasopharyngolaryngoscopy (VNL-1590STi, KayPentax, USA) and confirmed by laryngeal electromyography. All patients had left vocal fold paralysis for more than 6 months and had never undergone any medialisation procedure. UVFP was iatrogenic in 8 patients and idiopathic in 1 patient.

MRI Protocol

Imaging was performed on a 3T Philips Achieva TX MRI scanner (Philips Healthcare, The Netherlands). Sense Flex S coils of 11cm diameter were placed at the side of the neck covering the anterior part of the larynx below the angle of the mandible, while the participant lied supine on the scanner bed with the head resting on a headrest. The coils were held in place by putting a neck strap and sandbags at the side of the head and neck to minimize movement. The participants were asked keep the neck and jaw still and avoid throat clearing or swallowing during image

acquisition periods. The imaging protocol for DWI and DCE-MRI is shown in Table 6-1. DWI was always performed prior to DCE-MRI. For DCE-MRI, 0.2mmol/kg gadolinium, 3mls/second was injected at the beginning of the fourth acquisition of serial scan images.

Table 6-1: MRI protocol of DCE-MRI and Diffusion weighted imaging (DWI)

	DWI	DCE
Sequence Name	SE-EPI	3D T1-FFE
Repetition time (ms)	3749	4.5
Echo time (ms)	37	2.2
Image resolution	1.3x1.3	0.8x0.8
Slice thickness (mm)	5	3
Number of slices	14	24
Flip angle	90	10
Field of view (FOV)	100x53	150x150
Temporal resolution (seconds	NA	11.3
No of averages	6	1
Total acquisition time (minutes)	5 min 22 secs	6 min 36 secs
Dynamic scans	-	35
b-values	0, 50, 100	-

Image analysis

Images were downloaded to OsiriX (open-source software, Geneva, Switzerland) workstation for analysis (Md et al., 2004). Subsequently, an experienced otolaryngologist placed the region of interests (ROI).

The **DCE-MRI** 35 serial acquisitions were registered using Robust Data Decomposition Registration to minimise inter acquisition respiratory and pulsatile motion effects (Hamy et al., 2014). TA and PCA muscles were identified on DCE-MRI by an experienced radiologist unaware of the EMG results. Separate regions of

interests (ROI) were placed within the left and right TA and PCA muscles (Figure 6-1). Images affected by motion artefacts such as swallowing or throat clearing were excluded. A time intensity curves (TIC) were generated for the TA and PCA by plotting signal intensity against the time point for the 35 acquisitions (Figure 6-2). Signal intensity was normalised using the formula below (Langer et al., 2009):

$$C = {(S_t) - (S_{base})}\Big/{(r_1)\times(S_{base})}$$

by which S = signal intensity; t = time; C = concentration; and $r1$ = relaxation constant. The native T1 (tissue properties after injection) of the tissue before the injection of the contrast agent (T10) was assumed to be equal to 1second. Slope of initial enhancement (SoE), maximum enhancement (ME) and area under the curve (AUC) at 1 minute (AUC-1) were derived using the formulae in Table 6-2 (Horsthuis et al., 2009).

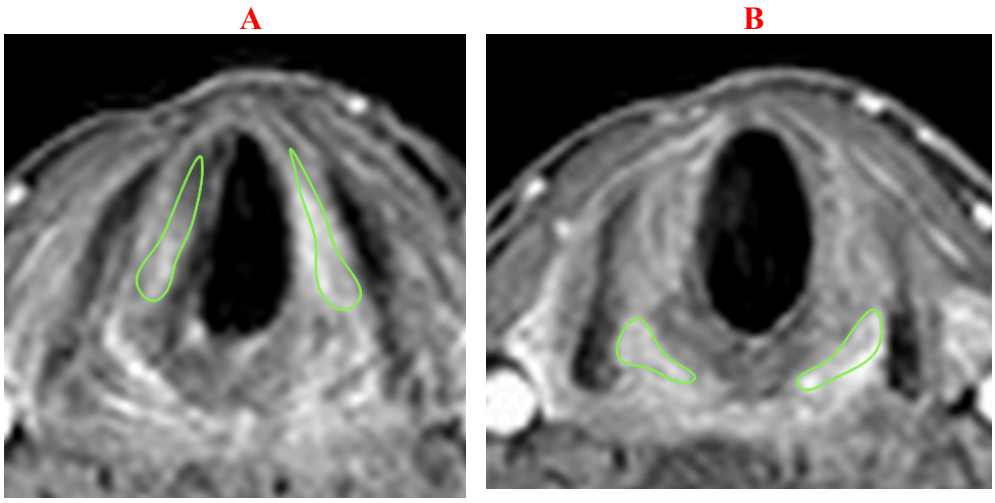


Figure 6-1: Post-contrast T1 weighted image from DCE MRI. The image demonstrates enhancement within the: A) Thyroarytenoid (TA) and B) Posterior cricoarytenoid (PCA) muscles. Regions of interest outlined in green lines within the TA and PCA are used to derive time versus signal intensity curve (below) from which quantitative metrics (SoE, ME, AUC-1) can be derived.

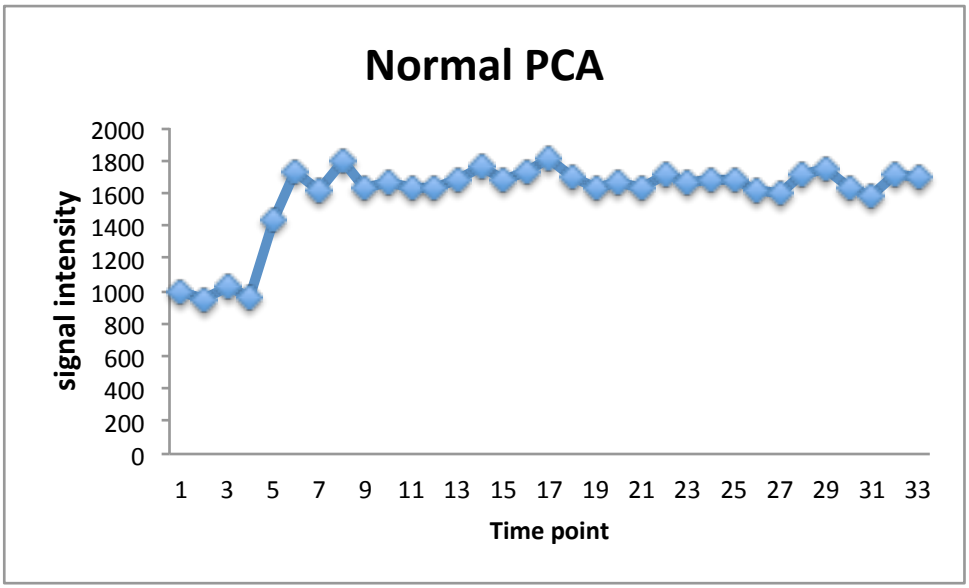


Figure 6-2: Time intensity curve (TIC)

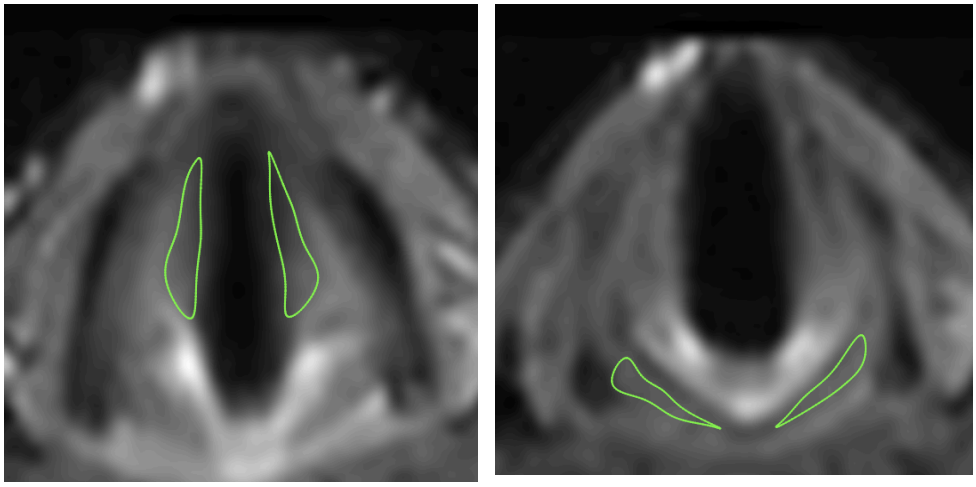


Figure 6-3: These are b0 images of diffusion weighted imaging of: A) TA; and B) PCA. The region of interests are drawn in green lines. 'b0 images' indicates images acquired at b-value 0 s/mm^2 in which the diffusion sensitising gradient is absent so free water in tissues appears bright due to the intrinsic T2-weighting.

Table 6-2: Formula for DCE-MRI metrics derived from the time intensity curve analysis

Metrics	Formula
Slope of enhancement	$(S_{t0+1}) - (S_{t0})$
Maximum enhancement	$(S_{peak}) - (S_{base}) / S_{base}$
Area under the curve up to time *t* (integral of TIC over the interval [0, *t*])	$\int_0^t TIC(t)dt$

S = signal intensity

t = time

C = concentration

$r1$ = relaxation constant

For **DWI,** separate ROIs were located on the b0 image within left and right TA and PCA muscles (Figure 6-3) and automatically propagated b50 and b100 images. ADC was calculated by performing a least square fit to the mean signal intensity of each muscle versus b-value using Mac Excel (2011) with solver (Punwani et al., 2013).

Statistical analysis

MRI derived metrics were compared:

- within volunteers, between left and right TA and PCA
- within patients, between paralysed and normal TA and PCA
- between volunteers and patients normal TA and PCA

- between volunteers normal and patients paralysed TA and PCA

Comparison between muscles within subjects was performed using the Wilcoxon signed rank test. The Mann Whitney test was used for comparison across volunteer and patient groups.

6.4 Results

A single patient's DCE-MRI dataset was excluded from the analysis due to severe residual motion artefact following RDDR registration. For DWI another patient was excluded from the analysis of PCA metrics due to severe image susceptibility related distortion precluding identification of the muscle.

Within volunteers

There was no significant difference for DCE and ADC metrics between left and right TA or PCA muscles (Table 6-3).

Table 6-3: This table shows within volunteers' median and interquartile ranges for DCE-MRI parameters and ADC of DWI of thyroarytenoid and posterior cricoarytenoid muscles.

Thyroarytenoid muscles

Metrics	Left TA (n = 5)	Right TA (n = 5)	Within volunteers (N vs N)
SoE	0.08 (0.05, 0.14)	0.06 (0.04, 0.12)	0.23
ME	0.56 (0.47, 1.00)	0.55 (0.48, 0.91)	0.34
AUC 1 minute	0.35 (0.33, 0.65)	0.34 (0.33, 0.65)	0.08
ADC ($x10^3$)	2.70 (1.90, 4.31)	2.87 (2.58, 4.05)	0.841

Posterior cricoarytenoid muscles

Metrics	Left PCA (n = 5)	Right PCA (n = 5)	Within volunteers (N vs N)
SoE	0.07 (0.05, 0.09)	0.06 (0.04, 0.12)	0.50
ME	0.59 (0.56, 0.62)	0.58 (0.54, 0.79)	0.50
AUC 1 minute	0.40 (0.38, 0.41)	0.42 (0.38, 0.45)	0.34
ADC ($x10^3$)	2.12 (1.72, 2.62)	2.42 (0.57)	0.034

The last column showed p value of Wilcoxon sign rank test.

Within patients

For TA**,** AUC1 and ADC were significantly lower on the paralysed side compared with the normal side (p=0.017, p=0.04); no significant difference was demonstrated for the remainder of the DCE metrics (p=0.16, p=0.09). For PCA, all DCE and ADC metrics were significantly lower on the paralysed side than the opposite normal (p=0.01) (Table 6-4).

Table 6-4: This table shows within patients' median and interquartile ranges for DCE-MRI parameters and ADC of DWI of thyroarytenoid and posterior cricoarytenoid muscles.

Thyroarytenoid muscles

Metrics	Paralysed (n = 8)	Normal (n = 8)	Within patients (paralysed vs N)
SoE	0.06 (0.04, 0.08)	0.07 (0.06, 0.09)	0.16
ME	0.50 (0.36, 0.64)	0.61 (0.48, 0.67)	0.09
AUC 1 minute	0.26 (0.25, 0.39)	0.36 (0.26, 0.46)	0.017*
ADC ($x10^3$)	1.68 (1.06, 2.28)	2.88 (2.32, 4.43)	0.04*

Posterior cricoarytenoid muscles

Metrics	Paralysed	Normal	Within patients (paralysed vs N)
SoE	0.06 (0.045, 0.07)	0.12 (0.11, 0.16)	0.01*
ME	0.43 (0.36, 0.52)	0.75 (0.69, 0.87)	0.01*
AUC 1 minute	0.30 (0.23, 0.37)	0.56 (0.48, 0.60)	0.01*
ADC ($x10^3$)	1.7 (1.49, 1.83)	2.63 (2.31, 3.51)	0.01*

The last column showed p value of Wilcoxon sign rank test.

Between volunteers and patients normal TA and PCA

For TA, there was no significant difference for SoE, ME, AUC-1 and ADC between volunteers and patients normal (p=0.5 to 1.0). For PCA, there was also no was significant difference for ADC (p=0.12) between volunteers and patients normal, but the SoE, ME and AUC-1 were significantly higher in the patients normal than volunteers normal (Table 6-5).

Table 6-5: This table shows results of comparison between volunteers and patients normal thyroarytenoid and posterior cricoarytenoid muscles.

Metrics	Patients versus volunteers normal TA	Patients versus volunteers normal PCA
SoE	1.00	0.001*
ME	0.96	0.03*
AUC 1 minute	0.52	0.01*
ADC ($x10^3$)	1.00	0.12

Data presented is p-values of Mann Whitney test.

Between volunteers and patients paralysed TA and PCA

For TA, ADC was significantly lower in the paralysed than the normal but DCE was not significantly different. For PCA, however, both DCE and ADC were significantly lower in the paralysed PCA and left and right PCA of volunteers, except for the SoE (Table 6-6).

Table 6-6: This table shows results of comparison between volunteers normal and patients paralysed thyroarytenoid posterior cricoarytenoid

Metrics	Paralysed vs L&R TA volunteers	Paralysed vs R&L PCA volunteers
SoE	0.32	0.08
ME	0.15	0.01*
AUC 1 minute	0.101	0.02*
ADC ($x10^3$)	0.008*	0.03*

Data presented is p-values of Mann Whitney test.

6.5 Discussion

Overview

Advances in MRI now allow the assessment of microstructure properties to differentiate normal and pathological tissues. Therefore, I explored the use of DCE-MRI and DWI to evaluate perfusion of TA and PCA in order to further understand the pathologic changes in human denervated laryngeal muscles.

Dynamic contrast enhanced MRI- DCE-MRI is a method of physiologic imaging that assesses tissue microcirculation, capillary permeability and composition of the interstitial space based on temporal signal intensity changes on T1-weighted imaging following extravasation of intravenous water-soluble contrast bolus injection to the interstitial space (Punwani, 2011). The initial SoE of the TIC relates to first pass tissue perfusion. It also relates to the capillary network, as a steep slope can indicate increased capillary density and permeability (Verstraete et al., 1996). Tissues with lower perfusion, capillary density and permeability have slower and sustained enhancement (Lavini et al., 2013). In contrary, increased capillary density or permeability due to neo-angiogenesis will result in rapid enhancement and rapid washout of contrast from tissue (Punwani, 2011). AUC a summation measure of contrast delivery and retention during the measurement period, is therefore reflective of changes in the capillary network (Punwani, 2011).

Diffusion weighted imaging- Application of motion probing magnetic field gradients to T2-weighted MRI enables measurement of molecular water motion including microcirculation of blood (Khoo et al., 2011). Signal loss is proportional to

the displacement of water molecules and the applied diffusion gradient (reflected by the b-value (Khoo et al., 2011). By acquiring a series of images with incremented b-values the displacement of water molecules can be quantified. At low diffusion weightings (b-values 0 to 100s/mm^2) most of the signal loss within an image occurs due to perfusion (Khoo et al., 2011; Takahara and Kwee, 2012; Yanagisawa et al., 2009a).

Synopsis of key findings

My main finding confirms paralysed laryngeal muscles have reduced tissue perfusion. We found ADC was significantly lower in paralysed TA and PCA compared with normal muscle. Additionally, perfusion parameters derived from DCE-MRI were also significantly lower in paralysed muscle. My second main observation was that the 'normal' PCA muscle within a patient with UVFP demonstrates higher tissue perfusion than normal PCA in healthy volunteers.

Comparison with previous work

Application of DWI in studies evaluating muscle perfusion using low b-values is limited. Yanagisawa et al. studied lower limb muscles perfusion following exercise. This study demonstrated that DWI is useful in evaluating perfusion and diffusion of intramuscular water molecules. It was found that muscle perfusion was increased after exercise and decreased with cooling. ADC at low b-values reflecting perfusion was confirmed when the reduction of ADC was demonstrated following temporal arterial occlusion (Yanagisawa et al., 2009a, 2009b). DWI in rabbits' denervated limb muscles was employed using high b-values to assess the diffusion of extracellular space. It demonstrated increased ADC values following acute skeletal

muscles denervation that may be due to an increase in local blood flow although histological examination depicted muscle atrophy without capillary enlargement (Goyault et al., 2012; Holl et al., 2008).

Reports on the use of DCE-MRI in studying the perfusion of skeletal muscle are scarce. Jiang et al and Taoka et al have evaluated perfusion within extra-ocular muscles in patients with Graves' disease using DCE-MRI (Jiang et al., 2012; Taoka et al., 2005). Both papers demonstrated a reduction in perfusion parameters that correlated with disease severity in patients with Graves disease. Histological changes of the extraocular muscles in patient with Graves disease include perivasculitis, oedema, fibrosis and abnormal vasculature (Zhao et al., 2011).

To my knowledge, the use of DCE-MRI and low b-values DWI in denervated muscles particularly laryngeal muscles has not been reported.

Studies on totally denervated limb skeletal muscles showed reduced capillary density and fibrosis (Borisov et al., 2000; Lu et al., 1997). In laryngeal paralysis, atrophy and fibrosis were demonstrated within 3 months of the onset of nerve transection but histological examination did not include measurement of capillary density (Shindo et al., 1992; Zealear et al., 1994). Microvascular density is known to be reduced in paralysed limb skeletal muscles 2 months after denervation (Borisov et al., 2000). To the best of my knowledge no publications have histologically evaluated microvascular change in denervated laryngeal muscles; however mitochondrial derangement has been observed in laryngeal muscles that could reflect abnormal perfusion (Leão et al., 2010). My results support histological evidence at different

skeletal muscle sites (Borisov et al., 2000; Lu et al., 1997) to confirm reduced perfusion in paralysed laryngeal muscles.

I also found that the normal PCA of patient group had increased perfusion as per DCE-MRI metrics. This may be due to overloading of the contralateral normal PCA to maintain the glottal airway. Eggington et al. demonstrated increased in capillary growth in rats skeletal muscles due to compensatory overload (Egginton et al., 1998). However, Lloyd et al. observed no further increase in capillary density after 12 days in overloaded rats muscles and no difference was found with the control group (Lloyd et al., 2003). Therefore further studies in this are is necessary for laryngeal muscles.

Limitation of the present study

Our study has several limitations. DWI is susceptible to artefacts particularly at tissue interfaces such as air to soft tissue for TA or bone to soft tissue for PCA causing geometric image distortion (Khoo et al., 2011) and this may have influenced my results. However I did not observe a systematic bias in the presence or absence of between paralysed and non-paralysed sides.

DCE-MRI also faces inherent motion artefacts susceptibility especially for TA due to respiration motion. One set of images of DWI and DCE-MRI was removed from analysis due to these issues. Respiratory motion artefacts may be the reason for the inability to demonstrate significant difference for DCE-MRI metrics between the paralysed and normal TA.

As commensurate with the nature of the patients included within this study I did not obtain histological correlates of MRI findings. Further pre-clinical work in this area remains to be performed.

The present study included a small sample of participants, limiting my inferences. However even with a small sample size I was able to demonstrate that assessment of tissue perfusion by MRI is feasible and significant differences in tissue perfusion parameters exist between paralysed and normal laryngeal muscles.

Clinical impact

DCE-MRI and DWI have the potential to be used to evaluate perfusion of denervated laryngeal muscles and may become useful tools to monitor physiologic changes following denervation and reinnervation. DWI may be preferable to DCE-MRI as it does not require intravenous contrast, thus increasing speed and decreasing risk and cost.

6.6 Conclusion

This study demonstrates the feasibility of assessing tissue perfusion in laryngeal muscles using DCE-MRI and DWI. The findings confirm that paralysed TA and PCA muscles have reduced perfusion and that apparently normal contralateral musculature in patients with UVFP has elevated perfusion. This work will inform mechanistic studies of vascular change in paralysed laryngeal muscles and clinical studies evaluating utility perfusion parameters in predicting and monitoring responses to treatment.

Chapter 7 Functional assessment of vocal fold mobility using cine-Magnetic Resonance Imaging

7.1 Abstract

Objective: To evaluate the repeatability of quantifying functional vocal fold mobility by cine magnetic resonance imaging (cine-MRI) in unilateral vocal fold paralysis (UVFP) patients.

Materials and methods: Cine-MRI of the vocal folds was performed on 5 healthy and 9 UVFP participants (gradient echo, temporal resolution: 0.7 seconds per three slices, axial plane). Imaging was performed during a 10-second period of quiet respiration and during phonation. Cine-MRI images were quantitatively analysed by measuring the area between the vocal fold and the midline during phonation (PA) and respiration (RA). Left to right ratio of PA (phonation ratio, PR) and RA (respiration ratio, RR) was calculated. Then the magnitude of asymmetry of left and right vocal fold during phonation (vocal fold phonation asymmetry, VFPa) and respiration (vocal fold respiration asymmetry, VFRa) was calculated as '1-PR' and '1-RR' respectively. In addition, the total glottal area was determined during phonation and respiration, and percentage reduction associated with phonation calculated (VFaP – vocal fold abduction potential). Intra-session, inter-scan repeatability of study metrics (VFPa, VFRa, VFaP) was assessed using the intraclass correlation coefficient (ICC). Mann Whitney test was used to compare VFPa, VFRa and VFaP differences between groups.

Results: ICC values of VFPa, VFRa and VFaP were 0.92, 0.95 and 0.90 respectively for intra-session, and 0.44, 0.94 and 0.82 respectively for inter-scan. VFPa ($p=0.012$), VFRa ($p=0.001$) were significantly lower and VFaP ($p=0.008$) were significantly higher for volunteers compared with patients.

Conclusion: Cine-MRI quantitative metrics show good repeatability (except VFPa) and feasibility for identification of patients with UVFP.

7.2 Introduction

A variety of techniques have been used to evaluate phonatory function of vocal folds qualitatively or quantitatively including videostroboscopy, videokymography and digital kymography to record vocal fold vibration (Omori et al., 1996; Rosen, 2005; Verikas et al., 2009). Videostroboscopy in particular is widely available and regarded as a vital assessment tool in voice clinics, where it may enhance the detection of laryngeal pathology such as vibration disorders that may be due to lesions, neurological problems and impaired mobility due to vocal fold paralysis.

In vocal fold paralysis, impaired abduction during respiration reduces the size of the glottal airway compromising breathing. Quantification of vocal fold mobility in terms abduction potential is necessary to provide a measure of effectiveness of interventions conducted to increase the glottal airway (Dailey et al., 2005). However, validated, non-invasive quantitative measures are lacking. Glottal area assessed during respiration by flexible endoscope can be used to quantify vocal fold abduction (Dailey et al., 2005; Motoyoshi et al., 2004; Waters et al., 1996), but accuracy is limited by distortion caused by varying distance from the object and the wide angle lens of the endoscope (Czaja Pawel et al., 2007; Dailey et al., 2005; Forkert et al., 1996; Motoyoshi et al., 2004; Waters et al., 1996). To correct for this, the insertion of a probe to measure the distance from the tip of scope to the true vocal fold has been suggested, but this increases the size of the endoscope required, is time-consuming and uncomfortable for routine clinical application (Brancatisano et al., 1983). Whilst endoscopic direct visualisation is possible it remains technically difficult as the camera must remain tangential to the plane of vocal folds with clear visualisation of the anterior commissure and vocal process is available (Dailey et al.,

2005). Imaging of the vocal folds with magnetic resonance imaging (MRI has previously been suggested as an alternative (Schlamann et al., 2009).

In this study, I aimed to develop and evaluate 3.0T MRI based quantitative measures of vocal fold mobility in normal volunteers and patients with feasibility for differentiating UVFP to determine feasibility for quantification of vocal fold mobility.

7.3 Methods

The local ethics committee approved this study and informed consent was obtained from all participants.

Study overview

This study investigates the reliability and reproducibility of the cine-MRI for quantitatively evaluating vocal fold mobility in UVFP alongside healthy controls. The cine-MRI derived parametrics were evaluated for:

1) within subject repeatability a) between repeated scans within the same session (intra-session) and b) between repeated scans in two separate sessions (inter-scan).

2) inter-reader reproducibility of parametricsbetween 2 separate observers.

3) application of quantitative parametrics for differentiation between healthy volunteers and UVFP patients.

Participant selection

Healthy controls: Five (2 female [age 39 and 40], 3 male [age 28, 31 and 35]) volunteers without any current or previous history of vocal fold conditions (excluding previous upper respiratory tract infection) were recruited among the local hospital staffs.

Unilateral vocal fold paralysis patients: 9 UVFP patients (7 female, 2 male) with age ranged from 19 to 62 years old were recruited from the local hospital ear nose and throat specialists voice clinics. Prior to recruitment, UVFP was identified by flexible nasopharyngolaryngoscope of (VNL-1590STi, KayPentax, USA) and confirmed by laryngeal electromyography. The causes were secondary to iatrogenic

or idiopathic causes. All patients identified had left vocal fold paralysis for more than 6 months and had never undergone any medialisation procedure. Patients who were claustrophobic or, had associated multiple lower cranial nerve palsies, vocal fold lesions, voice related neurological disorders (spasmodic dysphonia, voice tremor) or cricoarytenoid joint fixation were excluded from the study. Exclusion criteria also include presence of any implants in the neck region that may cause substantial image artifact.

Cine-MRI protocol

All participants were imaged on a 3T Philips Achieva TX MRI scanner (Philips Healthcare, The Netherlands). Sense Flex S coils of 11cm diameter were placed at the side of the neck covering the anterior part of the larynx below the angle of the mandible, while the participant lay supine on the scanner bed with the head resting on a headrest. The coils were held steady by placing a neck strap and sandbags at the side of the head and neck to minimize movement.

Details of the protocol used for the MRI are summarized in Table 7-1. Sagittal reference images were used to angle the axial cine-MRI slices parallel to the vocal folds. Subsequently, a T2-weighted turbo spin echo coronal acquisition was performed to cover the larynx. Three cine-MRI axial slices were then positioned to ensure that the vocal folds remained within the imaging volume accounting for the normal vertical elevation that occurs during phonation (Figure 7-1).

Table 7-1: MRI protocol for vocal fold motion and coronal T2-weighted imaging

	Vocal fold motion	T2-weighted
Sequence Name	Fast field echo (FFE)	SE-TSE
Repetition time (ms)	2.1	6789
Echo time (ms)	0.82	120
Image resolution	1.5x1.5	0.5x0.5
Slice thickness (mm)	10	2.5
Number of slices	3	14
Flip angle	20	90
Field of view (FOV) (mm)	240x180	140x140
Temporal resolution (seconds)	0.764	-
Total acquisition time	1 min 31 secs	5 min 32 secs
Number of dynamic scans	120	-

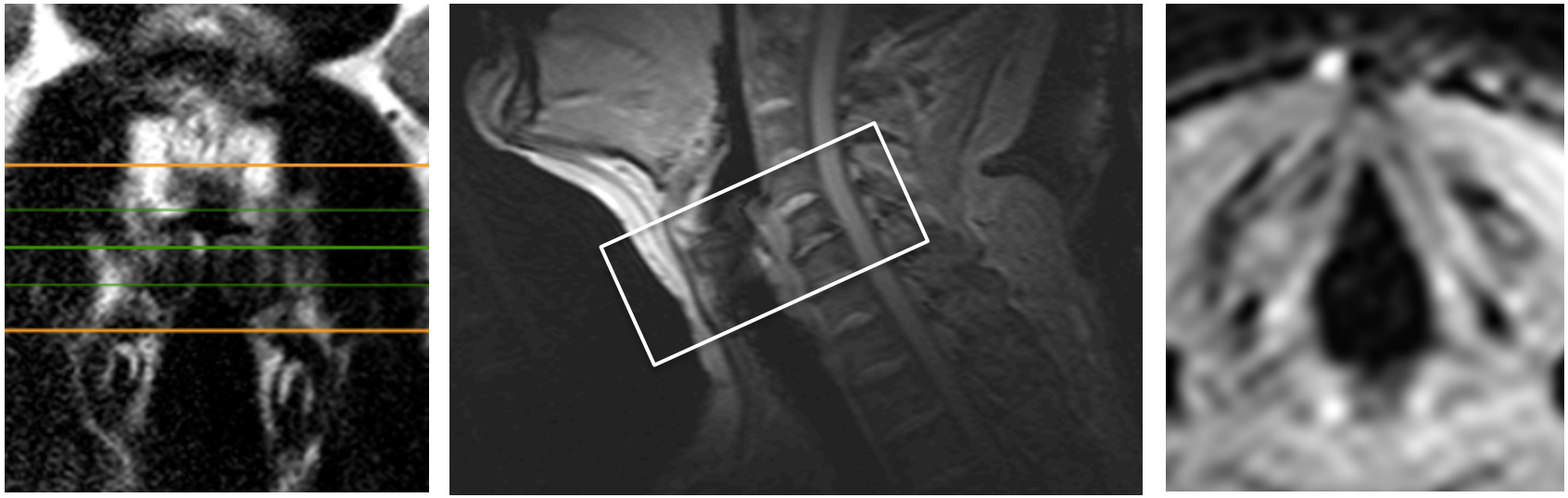


Figure 7-1: A) A T2-weighted coronal view of a participant showing three slices placement, two bold lines at the top and bottom and one line in the middle through the superior surface of the true vocal folds. B) A sagittal reference scan showing the angulation of the axial motion sequence positioned at right angles to the larynx. C) A frame from an axial motion sequence image of vocal folds during respiration.

Functional tasks for vocal fold assessment

Prior to imaging, participants were instructed in and practiced a standardised set of phonation tasks (Schlamann et al., 2009). During cine-MRI, participants were instructed to perform the standardized tasks. Specifically they were asked to phonate /hee/ for the first 10 seconds of scanning, then breathe quietly through the nose for 10 seconds; phonate /hee/ for a further 10 seconds; breath quietly again for 10 seconds and then phonate /hee/ for a final 10 seconds. Ten sequential cine-MRI acquisitions were performed during each /hee/ and quiet breathing period.

Definition of parametrics reflecting vocal fold mobility

Normal vocal folds open during respiration and close at the midline during phonation as shown in Figure 7-2. The glottal area is bounded by right and left vocal folds, anterior commissure and posterior commissure. For purposes of assessment, I assessed three metrics: vocal fold phonation asymmetry (VFPa), vocal fold respiration asymmetry (VFRa) and vocal fold abduction potential (VFaP).

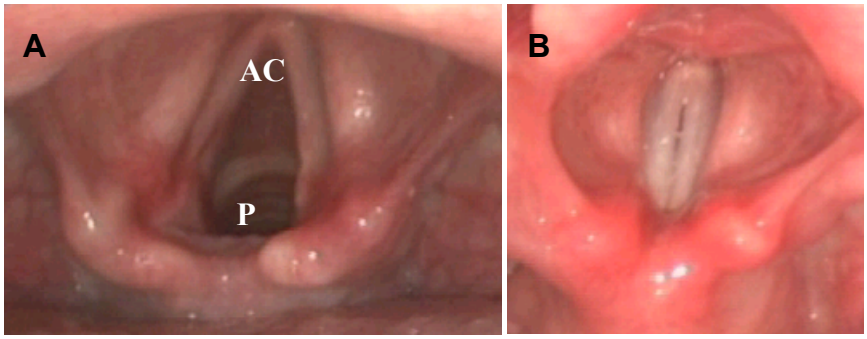


Figure 7-2: This endoscopic view of the normal larynx depicts the vocal folds position during: A) respiration; and B) phonation. AC: Anterior commissure, PC: Posterior commissure.

To derive the asymmetry indices I measured glottal area. Area measurements were performed separately for the left and right glottal areas using a bisecting line connecting the anterior commissure and posterior commissure (Figure 7-3). Phonation area (PA) and respiration area (RA) measurements were derived from images acquired during phonation and respiration respectively. The ratio between the left and right PA and and left and right RA was defined as the phonation ratio (PR) and respiration ratio (RR) respectively. We anticipated that PR and RR would be near to unity under normal conditions, reflecting the symmetrical movement of the right and left vocal folds. VFPa and VFRa were calculated as 1-PR and 1-RR respectively.

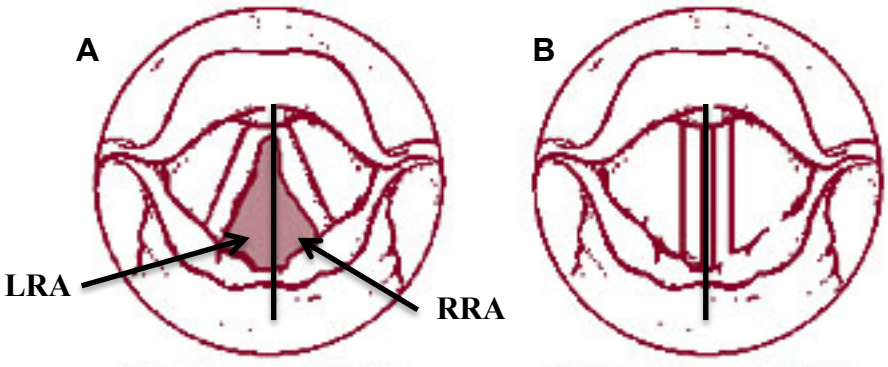


Figure 7-3: Schematic depicting an imaginary line bisecting the anterior and posterior commissure and dividing the: A) respiration; and B) phonation area into right and left. Note that the phonation area is very small. LRA: Left respiration area, RRA: Right respiration area.

This schematic diagram was taken from
http://en.wikipedia.org/wiki/File:Illu07_larynx02.jpg

In addition, the total glottal area during phonation and respiration was recorded and percentage reduction in the total glottal area associated with phonation calculated as VFaP. Paralysed vocal folds by definition are immobile and lie either in the lateral position or near to the midline during both phonation and respiration (Figure 7-4) (Woodson, 2007a).

The summary of the three derived metrics is shown in Table 7-2.

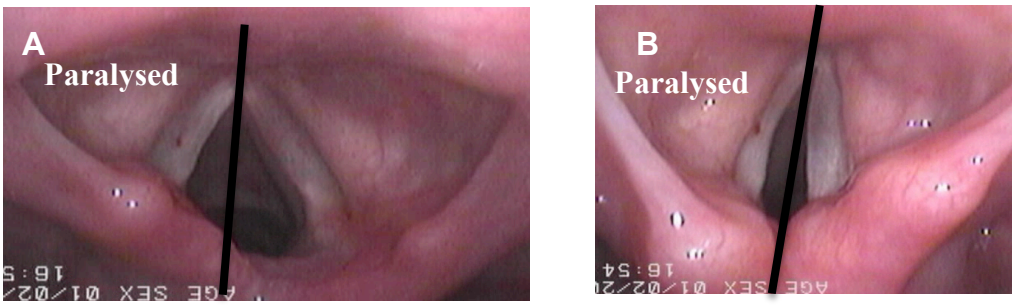


Figure 7-4: This endoscopic view of left vocal fold paralysis depicts the position of the paralysed vocal folds during respiration and phonation

Table 7-2: This table shows summary of vocal fold motion metrics and the inferences

Metrics	Definition and formula	Interpretation
PA	Glottal area during phonation. The PA is divided into left and right PA by an imaginary vertical line bisecting the anterior and posterior commissure	NA
RA	Glottal area during respiration. The RA is divided into left and right RA by an imaginary vertical line bisecting the anterior and posterior commissure	NA
VFPa	PR is ratio left to right PA = *LPA/RPA* Vocal fold ***phonation*** asymmetry (VFPa) = 1-PR	PR, close to one = normal symmetrical adduction. Therefore the 1-PR will be small PR, high deviation from one = asymmetrical adduction. Therefore the 1-PR will be large
VFRa	Ratio left to right RA = *LRA/RRA* Vocal fold ***respiration*** asymmetry (VFRa) = 1-RR	Close to one = normal symmetrical adduction. Therefore the 1-PR will be small High deviation from one = asymmetrical abduction. Therefore the 1-PR will be large
VFaP	Ability of the VFs to abduct by calculating the potential glottal area involved from adducted position (PA) to abducted position (RA) $[{RA - PA}/{RA}]\times 100$	Large VFaP = good potential glottal airway that may indicate both vocal folds mobile Small VFaP = small potential airway that may indicate one of the vocal folds is immobile

PA: phonation area
RA: respiration area
PR: Phonation ratio
RR: Respiration ratio
VFaP: Vocal fold abduction potential

Image analysis

Images were downloaded to OsiriX (open-source software, Geneva, Switzerland) (Md et al., 2004) workstation for analysis. The complete dataset for each participant comprised two intra-session cine-MRI studies (study 1 and study 2) and a third inter-scan cine-MRI study (study 3). Each study comprised 10 sequential cine-MRI acquisitions performed during phonation and a second set of 10 sequential cine-MRI acquisitions performed during respiration.

Two experienced readers independently performed measurements on each study. For each study the reader selected the axial cine-MRI slice that best represented glottal area for measurement (Figure 7-5). A vertical line was drawn from the anterior commissure bisecting the posterior commissure and the cervical vertebra deriving the triangular larynx to left and right glottal areas (Figure 7-5). A region of interest was then used to measure area between the bisecting vertical line medially and the vocal fold laterally and posteriorly (Figure 7-5). Area measurement was repeated for each sequential acquisition acquired during phonation and respiration respectively. Total glottal area was calculated by the sum of the left and right glottal area. An average phonation area and respiration area across the 10 acquisitions was derived and used for calculation of VFPa, VFRa and VFaP as defined earlier (Figure 7-6).

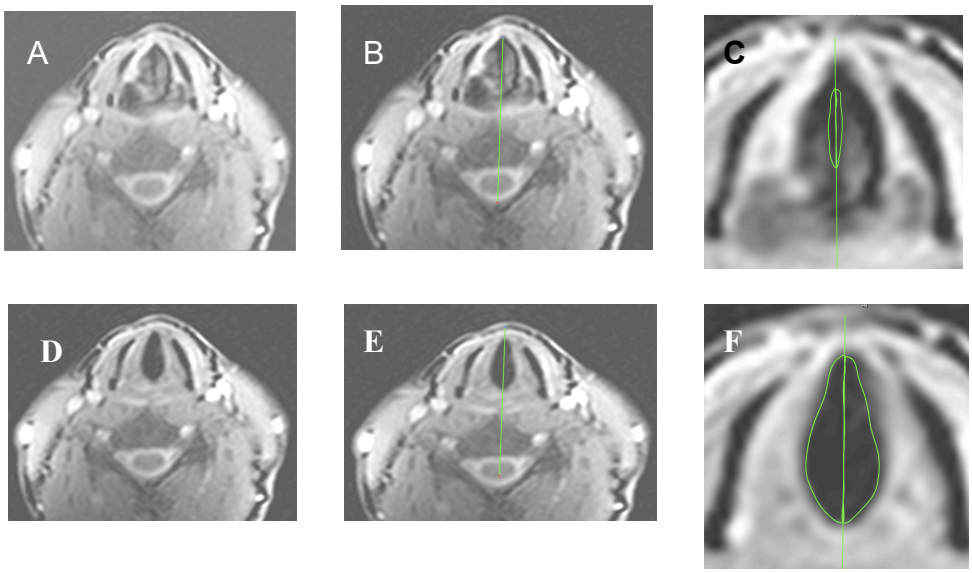


Figure 7-5: **A, D**: Selected image that represent glottal area during phonation and respiration respectively; **B, E**: A vertical line was drawn from anterior commissure bisecting the posterior commissure and the cervical vertebra dividing the triangular larynx to left and right glottal area; and **C, F**: Region of interests drawn to measure left and right phonation area and respiration area.

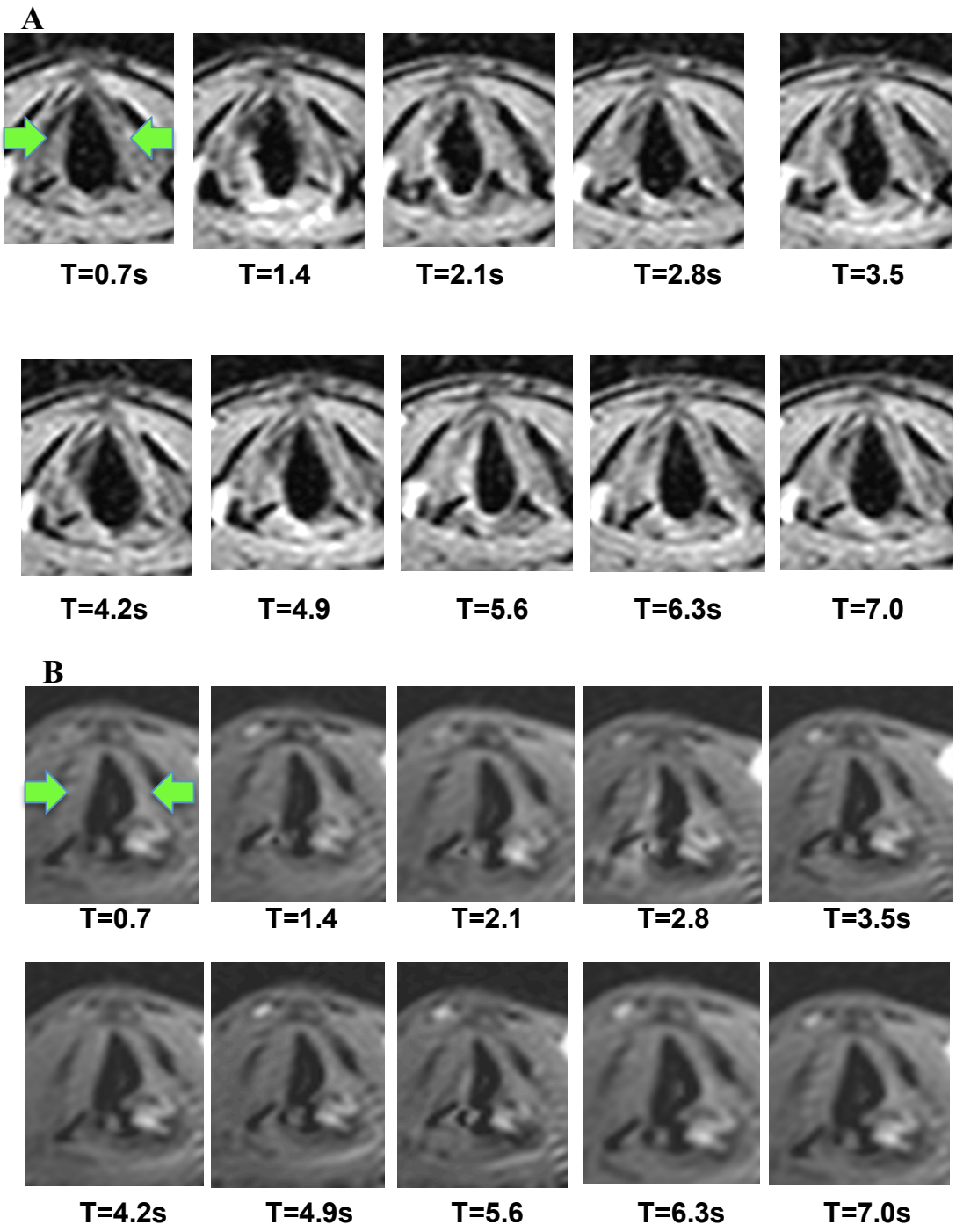


Figure 7-6: Selected sequential images of respiration at 10 time points of: A) a healthy volunteer; B) a patient with left vocal fold paralysis. The arrows shows right and left vocal fold

Statistical analysis

1) Within subject repeatability was tested using intraclass correlation coefficient test (ICC).

2) Inter-reader reproducibility in measuring the VFPa, VFRa and VFaP was assessed using ICC.

3) Application of quantitative parametrics for differentiation between healthy volunteers and UVFP patients was assessed by comparing the mean values of VFPa, VFRa and VFaP between groups using Mann Whitney test. Significance was defined at $p<0.05$.

ICC values ranged from 0 (not repeatable or reproducible) to 1 (perfect repeatability or reproducibility). The ICC has to be at least 0.6 for the measurements to be clinically useful (Chinn, 1991). Repeatability and reproducibility of the measurements was considered good for ICC of similar to 0.75 and moderate for ICC of between 0.5 and 0.75 (Portney and Watkins, 1993)

7.4 Results

The median and interquartile ranges (IQR) of VFPa, VFRa and VFaP for healthy volunteers was 0.07 (0.04, 0.13), 0.07 (0.02, 0.15) and 93.84 (76.22, 95.17)% respectively; and for patients they were 1.01 (0.35, 1.96), 0.38 (0.31, 0.71) and 59.69 (44.52, 71.78)% respectively (Figure 7-7). The median and IQR of all the metrics of the two intra-session cine-MRI studies (study 1 and study 2) and a third inter-scan cine-MRI study (study 3) are shown in Table 7-3.

Table 7-3: This table shows median and interquartile range of VFPa, VFRa and VFaP measurements of healthy volunteers and patients for measured for study-1, study-2 and study-3.

Metrics	Readers	Study 1 (first 10 sequential images of within the same scan)		Study 2 (second sequential10 images of within the same scan)		Study 3 (the first 10 sequential images of the second visit scan)	
		Volunteer (n = 5)	Patient (n =9)	Volunteer (n = 5)	Patient (n =9)	Volunteer (n = 5)	Patient (n =5)
VFPa	Reader1	0.07 (0.04, 0.13)	1.01 (0.35, 1.96)	0.13 (0.04, 0.18)	0.96 (0.32, 2.30)	0.14 (0.06, 0.29)	0.37 (0.10, 0.78)
	Reader2	0.27 (0.06, 0.37)	1.0 (0.26, 1.50)	ND	ND	0.14 (0.11, 0.29)	0.37 (0.16, 0.48)
VFRa	Reader1	0.07 (0.02, 0.15)	0.38 (0.31, 0.71)	0.03 (0.01, 0.12)	0.43 (0.35, 0.71)	0.06 (0.02, 0.19)	0.73 (0.67, 0.80)
	Reader2	0.12 (0.08, 0.26)	0.63 (0.43, 0.72)	ND	ND	0.06 (0.05, 0.12)	0.73 (0.72, 0.80)
VFaP	Reader1	93.84 (76.22, 95.17)	59.69 (44.52, 71.78)	89.28 (67.74, 93.82)	53.82 (67.74, 93.82)	93.26 (88.05, 95.67)	73.08 (60.03, 85.40)
	Reader2	90.15 (70.09, 93.21)	52.92 (32.04, 71.98)	ND	ND	93.25 (93.01, 95.36)	73.08 (60.07, 83.49)

ND = not done

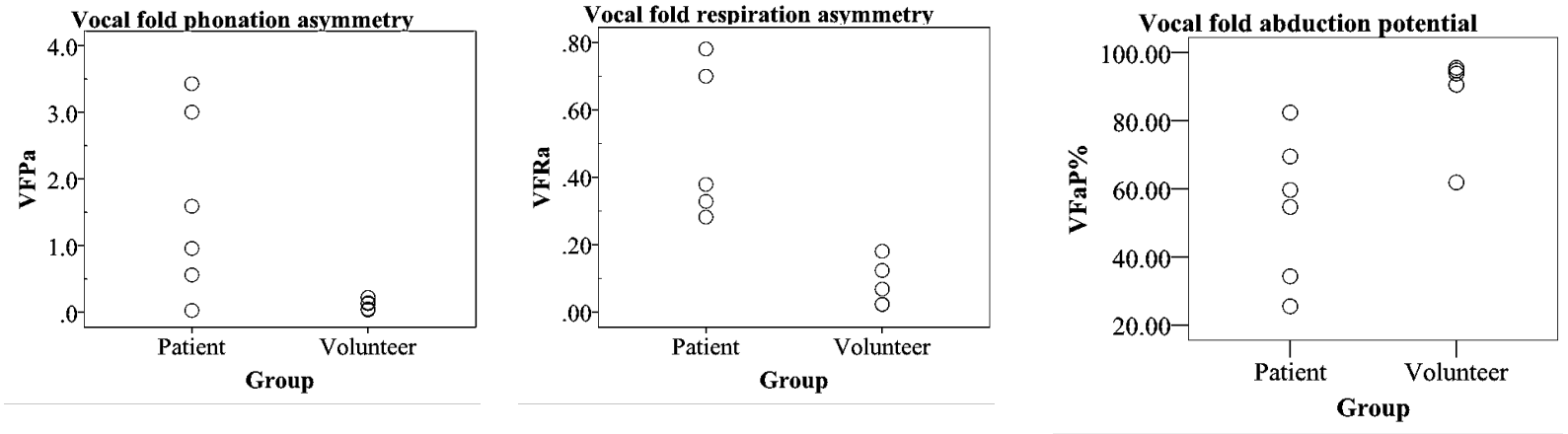


Figure 7-7 These graphs show distribution of VFPa, VFRa and VFaP in left vocal fold paralysis patients and healthy volunteers.

Results for within subject repeatability and inter-reader reproducibility are summarised in Table 7-4**.**

Within subject repeatability - Intra-session and Inter-scan variation

Intra-session ICC values and the 95% confidence interval (CI) of VFPa, VFRa and VFaP of both volunteers and patients was 0.92 (0.76, 0.97), 0.95 (0.85, 0.98) and 0.90 (0.69, 0.97) respectively. Inter-scan ICC values (95% CI) of VFPa, VFRa and VFaP was 0.44 (-0.24, 0.83), 0.94 (0.78, 0.98) and 0.82 (0.62, 0.97).

Inter-reader reproducibility

The ICC values (95% CI) of inter-reader reproducibility in measuring the VFPa, VFRa and VFaP are 0.84 (0.66, 0.93), 0.82 (0.60, 0.92) and 0.91 (0.77, 0.96) respectively.

Table 7-4: Results of intra-session, inter-scan repeatability, and inter-reader reproducibility.

Metrics	Intra-session	Inter-scan	Inter-reader
VFPa	0.92 (0.76, 0.97)	0.44 (-0.24, 0.83)	0.84 (0.66, 0.93)
VFRa	0.95 (0.85, 0.98)	0.94 (0.78, 0.98)	0.82 (0.60, 0.92)
VFaP	0.90 (0.69, 0.97)	0.82 (0.62, 0.97)	0.95 (0.88, 0.98)

Results presented are ICC values and the 95% confidence interval in the bracket

Comparison between healthy volunteers and patients

There was a significant difference in VFPa ($p=0.012$), VFRa ($p=0.001$) and VFaP ($p=0.008$) between volunteers and patients. The VFPa and VFRa were greater in patients than in volunteers indicating asymmetrical mobility of the vocal folds during phonation and breathing. The VFaP in the volunteers demonstrated a larger functional glottal airway during breathing than in patients, as both normal vocal folds in volunteers abduct symmetrically.

7.5 Discussion

Overview

Objective measurement of glottal airway is important in laryngeal research to measure whether an intervention improves vocal fold mobility. Impaired vocal fold mobility leads to glottal airway restriction that causes breathing difficulty (Dailey et al., 2005; Marina et al., 2011; Waters et al., 1996). To date, non-invasive objective measurements of vocal fold mobility are lacking. Cine MRI enables non-invasive acquisition of vocal fold adduction and abduction (Ahmad et al., 2006, 2009; Amin et al., 2012; Breyer et al., 2009; Davis et al., 1996; Faust et al., 2001; Schlamann et al., 2009; Tian et al., 2010). We further developed and assessed this technique for quantitative assessment of vocal fold mobility.

Synopsis of key findings

Theoretically, symmetrical mobility of vocal folds during phonation and respiration, will result in PR and RR near to ratio of one. Magnitude asymmetry of vocal folds area during phonation and respiration, VFPa and VFRa, was calculated to determine how deviated is the ratio from unity. Findings in the present study demonstrated that the magnitude asymmetry was small in volunteers and large in patients. VFaP in volunteers was larger than patients indicating that both vocal folds abduct producing larger glottal airway. There was significant difference demonstrated for VFPa, VFRa and VFaP between volunteers and patients. This shows that cine-MRI can be used to differentiate normal and abnormal vocal fold mobility.

Furthermore, the present study demonstrates high repeatability of the cine-MRI technique for measuring vocal fold mobility metrics VFRa and VFaP (ICC > 0.89). Inter-reader reproducibility was also almost perfect (ICC > 0.81).

Comparison with previous works

Previous work in a small number of paediatric patients with airway disorders has shown that cine-MRI of the larynx during phonation and respiration correlates with endoscopy findings (Faust et al., 2001); only one UVFP patient was included in this cohort. Schlamann et al studied 12 patients with hoarseness who were scanned using the cine-MRI technique and radiologists were able to qualitatively assess images to correctly diagnose vocal fold mobility disorders (Schlamann et al., 2009). To the best of my knowledge, only one previous study has measured quantitative mobility of the vocal folds (Ahmad et al., 2009). Ahmad et al. studied quantified horizontal vocal fold displacement during phonation in healthy volunteers (Ahmad et al., 2009) on cine-MRI by measuring change in the angle of the anterior commissure. Our study systematically assessed the utility of a range of cine-MRI derivable quantitative measures of vocal fold mobility, specifically derived from measurements of changes in the glottal area between the vocal folds during phonation and respiration. Our study is the first to assess the utility of the defined quantitative metrics in measuring vocal fold mobility during phonation as well as respiration in healthy volunteers and patients, and investigate the repeatability and reproducibility metrics.

Of note, I do not expect that the temporal resolution of dynamic scans within this or previous studies is sufficient to capture the vibratory motion of the vocal fold; however it is an attempt to capture the abduction/adduction potential of the vocal

folds. The temporal resolution of my study is comparable to the study published by Schlamann et al. (Schlamann et al., 2009). We are unable to compare the temporal resolution of my dynamic acquisition with the other two studies, as this has not explicitly been described within the respective manuscripts (Ahmad et al., 2009; Faust et al., 2001). A summary of cine-MRI protocols comparing my study with others is presented in Table 7-5.

Table 7-5: Comparison between studies that investigate vocal fold mobility.

RapidMRI protocol	Present study	Ahmad et al. 2009	Faust et al 2001	Schlamann et al. 2009
Participants	Healthy volunteers and UVFP patients	Healthy volunteers	Paediatric patients with airway disorder	Patients with hoarseness that may or not had vocal fold palsy
MRI machine	3.0T	1.5T	1.5T	1.5T
Sequence Name	Fast field echo (FFE)	T1-weighted spin echo sequence	turboFLASH	Steady-state free precession sequence (SSFP)
Planes	axial	axial	Axial and coronal	coronal
Repetition time (ms)	1.98	128.0	2.5	282
Echo time (ms)	0.78	4.0	1.2	1.3
Image resolution	160x120	128x128	128x128	128x128
Slice thickness(mm)	10	5	8	4
Number of slices	3	3	Not mentioned	2
Flip angle(°)	20	90	10	55
Field of view (FOV)	240mmx180mm	220mm	300cm	235x117.5mm^2
Temporal resolution (s)	0.736s (3 images per 0.734s)	Not mentioned	Not mentioned	4 images per second over 14s
Total acquisition time	1 minute, 31 seconds	36.9 seconds	10 seconds per slice (32 or 64 images)	14 seconds

Limitation of the present study

My study has some limitations. Paralysed vocal folds have shorter length and at lower vertical level than the normal opposites side. I found the true edge of the vocal fold was sometimes not visible on the acquired cine-MRI. Previous studies have shown that during phonation, normal larynx elevates (Ahmad et al., 2009) and paralysed vocal folds lack of the vertical elevation (Faust et al., 2001). This maybe a contributing factor for the poor inter-scan repeatability for VFPa measurements. However, other measures in the present study, VFRa and VFaP demonstrated high repeatability and appear unaffected by this limitation. Furthermore to minimise potential errors caused by the elevation of the larynx, I acquired three axial slices to capture the level which best represents glottal area during phonation; where elevation occurred I selected the slice that best represented glottal area for quantitative analysis.

Although limited by a small number of participants, the present study confirms the feasibility of quantifying vocal fold mobility using cine-MRI, and indicates capability of cine-MRI measures to reliably identify UVFP.

Clinical impact

I expect this technique would be of great benefit to non-invasively and objectively assess the graded improvement in vocal fold mobility that are likely to occur following surgical interventions aimed at re-establishing the dynamic movement of vocal fold.

7.6 Conclusion

I have shown that cine-MRI can objectively quantify vocal fold mobility and detect UVFP by derived quantitative metrics. The technique can be utilised within clinical trials of surgical interventions as a non-invasive measure of therapeutic success.

Chapter 8 Case series of non-selective and selective laryngeal reinnervation

8.1 Non-selective laryngeal reinnervation in unilateral vocal fold paralysis patients

8.1.1 Abstract

Objective: There are few reported series of non-selective laryngeal reinnervation with well-validated subjective, objective and radiological outcome measures, and so speculation remains as to the effectiveness of this intervention for patients with unilateral vocal fold paralysis (UVFP). I present a small case series with a range of such outcome measures collected prospectively.

Methods: Five patients with UVFP (3 females, 42 to 51 years old; 2 males, 32 and 60 years old) with range of 5 to 24 months duration of vocal fold paralysis underwent non-selective laryngeal reinnervation using ansa cervicalis to recurrent laryngeal nerve technique with concomitant collagen injection laryngoplasty. All of them were subjected to multidimensional outcome measures (voice handicap index-10: VHI-10, voice perceptual evaluation, acoustic analysis, maximum phonation time, aerodynamic analysis and video-laryngostroboscopy analysis) in which VHI-10 was the primary outcome measure, at baseline, 3-, 6-, and 12-month post-reinnervation. Laryngeal electromyography (LEMG) and T2-weighted MRI of the larynx was performed at baseline and at12-months to measure the neuromuscular integrity of the thyroarytenoid muscle (TA). Three of five patients completed a 12-month review.

Results: Voice improvement was achieved in all patients in whom the mean and standard deviation (SD) of VHI-10 scores were of 26.6 (4.83), 14.8 (8.53) and 7.8 (7.46) at baseline, 3- and 6-month, respectively. The normal score was reached at 6-months and maintained at 12-months in 2 of 3 patients (score: 6 and 7) who had completed the 12-month review. The motor units recruitment on LEMG was ‘very

much better' when compared to the baseline. There was higher signal intensity changes detected on the paralysed TA compared to the opposite 'control' on T2-MRI images at baseline. Repeat T2-MRI at 12-month showed normalisation of the signal intensity compared to the baseline values. One of the 3 patients suffered a neck abscess post-operatively, with a VHI-10 score of 21 at 12-months, only slight subjective improvement in LEMG motor unit recruitment and sustained high signal T2-MRI at 12-months.

Conclusion: A case series of 5 patients undergoing non-selective laryngeal reinnervation using an ansa-RLN technique with concomitant injection laryngoplasty is presented. Voice improvement was demonstrated by VHI-10 and other multidimensional outcome measures, and these were supported by LEMG and T2-MRI outcomes. To my knowledge, this is the first multidimensional prospective study of laryngeal reinnervation and also the first to suggest that 3T MRI may be a promising outcome measure for future reinnervation trials.

8.1.2 Background

Unilateral vocal fold paralysis (UVFP) occurs may be due to injury to the vagus nerve or, more commonly, its branch, the recurrent laryngeal nerve. The affected vocal fold become flaccid and immobile causing dysphonia and may be aspiration. It loses properties that determine the quality of voice like the mass, elasticity, resistance and symmetry of the vocal folds (Zheng et al., 1996). Different treatment options are available for the management of UVFP which include injection laryngoplasty, type I thyroplasty, arytenoid adduction and laryngeal reinnervation (Crumley, 1991; Isshiki et al., 1975, 1978; McCulloch and Hoffman, 1998)

Non-selective laryngeal reinnervation, refers to restoration of innervation to the larynx by either anastomosis or implant technique using suitable donor nerves (Crumley, 1991; Paniello, 2000; Tucker and Rusnov, 1981). It is a safe surgery (Blumin and Merati, 2008) that aims to re-establish the tone and bulk of the vocal folds by achieving "favourable laryngeal synkinesis" (Crumley, 2000). The most common technique reported is ansa to recurrent laryngeal nerve anastomosis (ansa-RLN) (Aynehchi et al., 2010). Non-selective laryngeal reinnervation using ansa-RLN anastomosis was re-introduced by Crumley and subsequently there were reports on the voice outcomes of this technique in UVFP (Crumley, 1991; Lee et al., 2007; Lorenz et al., 2008; Olson et al., 1998; Paniello et al., 2011; Smith et al., 2008; Wang et al., 2011; Zheng et al., 1996). These reported improvement of voice perceptual evaluation, video-laryngostroboscopy, laryngeal electromyography (LEMG) and acoustic analysis to varying degrees (Aynehchi et al., 2010).

Non-selective laryngeal reinnervation has the potential of restoring normal or near normal voice in UVFP patients. There are several advantages of laryngeal reinnervation over other surgical techniques. Specifically, it avoids vocal fold atrophy and provides a long-term solution by maintaining the vocal fold mass, bulk, tone and pliability for normal mucosal waves (Lorenz et al., 2008; Wang et al., 2011; Zheng et al., 1996). The possibility of other forms of laryngoplasty is preserved should voice outcomes reach desired levels. However, voice improvement takes up to 4 months (Lee et al., 2007; Lorenz et al., 2008; Maronian et al., 2003) . Concomitant injection laryngoplasty using absorbable material may strengthen the voice during this waiting period (Lee et al., 2007; Lorenz et al., 2008; Olson et al., 1998).

Here, the results of non-selective reinnervation with concomitant collagen injection laryngoplasty in five patients are presented. The primary voice outcome is voice handicap index-10 (VHI-10). The correlation between VHI-10 and neurological status as shown on LEMG and MRI is reported. Results of other multidimensional outcome measures are shown.

8.1.3 Methods

Local ethics committee approval was obtained for the present study.

Participant selection

The study population included 5 UVFP patients (3 female, 42 to 51 years old ; 2 male, 32 and 60 years old) who were diagnosed UVFP and met inclusion and exclusion criteria listed in chapter 2. They were recruited from the Royal National Throat Nose Ear Hospital and St. Thomas' and Guy's Hospital London in 2011-2012. The causes of the UVFP were thyroidectomy (2 patients), thymectomy (1 patient), vagal schwannoma excision (1 patient) and cervical spine procedure (1 patient). Duration of palsy was ranged from 5 months to 24 months.

Surgical procedure

For description of surgical technique, please see 1.6, page 41 to 42.

Temporary medialisation of the vocal fold was done at the same setting by injecting porcine collagen (Permacol™, Tissue Science Laboratories plc, "TSL" Aldershot UK) endoscopically

Outcome measures

The patients were assessed pre- and post-operatively for voice improvement using multidimensional measures and effect of reinnervation on the thyroarytenoid muscles as the following:

a) voice handicap index-10 (VHI-10) (Rosen et al., 2004);
b) maximum phonation time, seconds (MPT);
c) voice perceptual evaluation (GRBAS) (Hirano, 1981);
d) acoustic analysis (jitter %, shimmer %, NHR);
e) aerodynamic analysis (maximum sound pressure level (SPL) dB, and mean airflow rate (MFR) at voicing L/s);
f) video-laryngostroboscopy;
g) Laryngeal electromyography (LEMG); and
h) T2-weighted MRI (T2-MRI) of larynx

The primary outcome measure was VHI-10. Outcome measures for a) to f) above were performed at baseline and at 3-, 6- and 12-months after operation whereas LEMG and T2-MRI for neurological status assessment were done at baseline and 12-months after operation. To date, 3 patients have reached 12 months' assessment whereas the other 2 patients have reached 6 months' assessment.

Description of outcome measures, randomisation and blinding

Please see chapter 1, page 68 to 71 for the more detail description of outcome measures.

1. Voice Handicap Index 10 (VHI 10)

   The patients answered the questions independently and the total score was obtained.

2. Maximum phonation time (MPT)

   MPT is a measurement of maximum vocal performance that was assessed manually using a stopwatch to time the duration. The patient was asked to sustain /a/ at a comfortable loudness of voice after taking a normal breathe. This task was repeated three times and the longest duration was recorded.

3. Voice perceptual evaluation

   Audio files of patients reading the 'rainbow passage' were rated by 3 raters who have established high inter-rater reliability (Webb et al., 2003). The audio files were anonymised and randomised. The raters were unaware whether the audio files were recorded pre-operatively or post-operatively.

4. Acoustic analysis

   Voice recordings were performed in a quiet room using Multidimensional Voice Program (MDVP, KayPentax, NJ, USA). The microphone provided for the MDVP was placed at 45 degree from the angle of the mouth with 5 cm lips-to-microphone distance. The recording was done while the patient sitting on a chair. Participants were instructed to say vowel /a/ for 6 seconds at comfortable

loudness. A five second recording of a sustained vowel /a/ was used to measure F0, jitter percent, shimmer percent and NHR. The first 500ms of the start and the end of the vowel was removed t o avoid the start and end effect.

5. Aerodynamic analysis

Aerodynamic analysis using Aerophone (model 6600, Kaypentax, Lincoln Park USA) was done to measure the air leakage during phonation that indirectly measures closure of glottic gap. Parameters measured were- maximum SPL dB, and MFR at voicing L/s that were indirectly calculated from intra-oral articulatory pressures. Each patient was fitted with a standard anaesthesia face-mask covering the nose and mouth. A small tube that connects the machine and the patient's mouth was placed just slightly beyond the upper and lower incisors. The patient was then asked to say /pa//pa//pa/ for 3 times at minimum, comfortable and loudest loudness level. There was no phonation break at each loudness level once data acquisition was initiated. The device was connected to a computer system with software that generate the data required.

6. Videolaryngostroboscopic analysis

Please see chapter 4 as well.

7. Laryngeal electromyography (LEMG)

With patients supine and with their necks slightly extended, an experienced laryngologist and electromyographer performed LEMG at rest and during phonation using a monopolar needle inserted percutaneously into the TA muscles. Tracings were converted to executable files (exe.file) in a random

order, and then rated by a clinical neurophysiologist (SF) blinded to time-point, who evaluated whether (a) one of the tracings demonstrated slightly better recruitment and contained more normal motor units; (b) one of the tracings demonstrated very much better recruitment and contained more normal motor units; or (c) there was no difference between the 2 tracings.

8. T2-weighted magnetic resonance imaging (T2-MRI)

   Please see chapter 5 for the details of T2-MRI protocol and image analysis. In the present study, TA signal intensity was measured from the acquired images.

Statistical analysis

Data of the present study was presented descriptively using mean and standard deviation (SD) as well as in graphs for each measurement for individual patient.

8.1.4 Results

Results of 5 patients at baseline, 3-, 6- and 12-month are presented. Three patients (patient-1,-2 and -3) had completed the 12-month review. Two patients experienced complications: 1) Patient-2 had a neck abscess 3 weeks after the laryngeal reinnervation requiring drainage and antibiotics; and 2) Patient-4 had temporary dysphagia requiring extended hospital stay (10 days). This resolved completely.

Table 8-1: This table shows the patients' demographics, and VHI-10 scores, MPT and laryngeal electromyography motor unit recruitment

	Patient-1	Patient-2	Patient-3	Patient-4	Patient-5
Age	40	62	32	42	56
Gender	F	M	M	F	F
Palsy side	Left	Left	Left	Right	Left
Duration of palsy	5 months	2 years	1 year	1 year	1 year
Aetiology	Thyr	Vagal para-ganglioma excision	Thyr	Cervical spine procedure (anterior approach)	Thymoma excision
VHI-10	21	32	29	29	22
MPT	3	8	7	4	7.6
LEMG – motor unit recruitment	1	1	1	1	2

Thyr: thyroidectomy; MPT: maximum phonation time

Voice handicap index-10, laryngeal electromyography (LEMG) and T2-MRI

A maximum score for VHI-10 is 40, indicating a severe impact to the patient's physical, functional and emotional concern. Voice improvement was observed in all patients in whom the scores were improved at 3-, 6- and 12-month from the baseline at (Table 8-2). The normal score for the VHI-10 was reached at 6-month and it was maintained at 12-month in 2 of 3 patients who had completed the 12-month review. For patient-2 who had the neck abscess and completed the 12-month review, the VHI-10 had improved from 32 to 21 at 12-month, but it did not surpass the score of 11, which is the normal cut off (Arffa et al., 2012).

Table 8-2: VHI-10 scores at pre- and post-operative

Patient	Baseline	3-month	6-month	12-month
Patient-1	21	9	6	1
Patient-2	32	17	21	21
Patient-3	29	28	4	7
Patient-4	29	14	5	-
Patient-5	22	6	3	-

LEMG tracings of the 3 patients who had completed the 12-month review revealed that the motor units recruitment is very much better for patient-1 and -3, and slightly better for patient-2 when compared with the baseline.

There was higher signal intensity changes detected on the paralysed (left) thyroarytenoid muscles (TA) compared to the opposite normal on the T2-MRI images at baseline (Figure 8-1). Repeat T2-MRI at 12-month showed a reduction in the signal intensity of the paralysed TA compared to the baseline values (Table 8-3). For the normal TA, the signals had remained similar. However, for patient-2, the VHI-10 score was still abnormal at 12-month, the LEMG motor unit recruitment was

only slightly better and the high signal changes on the T2-MRI had remained at 12-months.

Normalisation of signal intensity in the 2 patients (patient-1 and -3) at the 12-month review demonstrated that the voice improvement at this stage was not due to the collagen injection but due to the reinnervation, since collagen is reabsorbed at between 4 and 6 months (Mallur and Rosen, 2010).

Table 8-3: Thyroarytenoid muscles signal intensity changes at pre- and post-operation and LEMG motor unit recruitment evaluation comparing pre and post-operative tracings

	LTA SI (baseline)	LTA SI (12-month)	RTA SI (baseline)	RTA SI (12-month)	Motor unit recruitment
Patient-1	4.00	1.88	1.82	1.71	very much better
Patient-2	3.47	3.48	2.48	2.20	slightly better
Patient-3	3.03	1.85	1.51	1.72	very much better

L, left; TA, thyroarytenoid; and SI, signal intensity

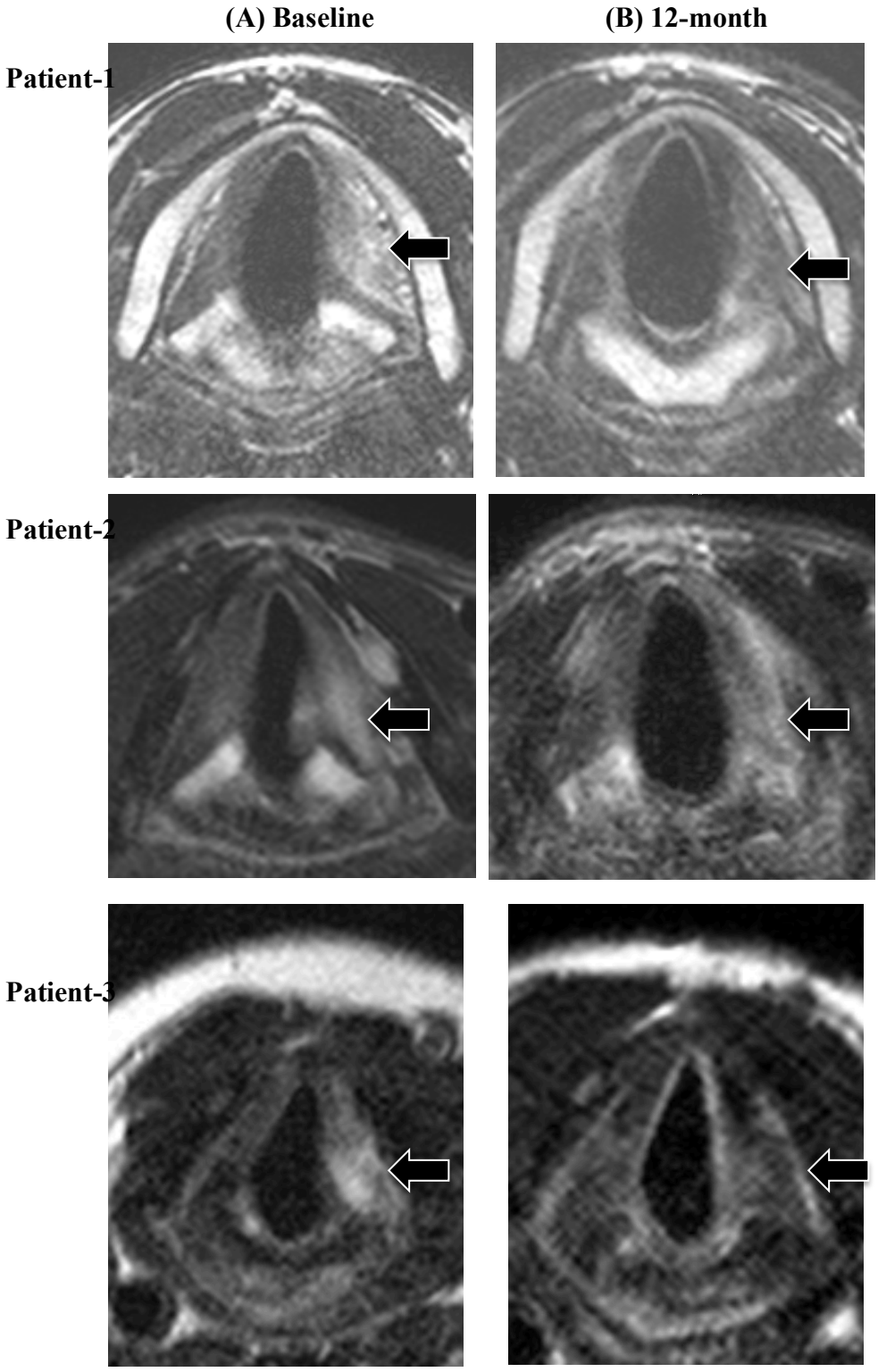


Figure 8-1: T2-weighted imaging of vocal folds images of patient-1, -2 and -3. Images on column (A) are taken at baseline and column (B) at 12-month. At baseline, the paralysed thyroarytenoid muscles (white arrow) had higher signal intensity than the opposite normal. The signal intensity normalised at 12-month except in patient-2

Maximum phonation time (MPT) and aerodynamic analysis

The patients were able to sustain phonation longer following reinnervation: MPT at baseline, 3-, 6-month was 5.92 (2.26)s, 9.20 (4.09)s and 11.6 (4.10)s. At 12-month, the MPT of the 3 patients reaching this time-point were of 18, 17 and 17. The patient who experienced postoperative abscess reported the worst VHI-10 score, and had an MPT within normal range (17). Figure 8-2 shows MPT at different time points for each patient.

For aerodynamic analysis, the maximum SPL dB and MFR L/s at baseline, 3-, 6-month were 74.46 (1.60) dB, 76.74 (2.86) dB and 76.92 (2.61) dB respectively for maximum SPL; and 0.46 (0.16) L/s, 0.30 (0.11) L/s and 0.26 (0.10) L/s respectively for MFR. At 12-months, the maximum SPL and MFR of the 3 patients who had completed follow-up were 90.62, 85.74 and 87.41 dB respectively for maximum SPL; and 0.25, 0.39 and 0.29 L/s respectively for MFR. Figure 8-2 shows the maximum SPL and MFR for each patient at different time points. The intensity is louder and air leakage lower than pre-operatively.

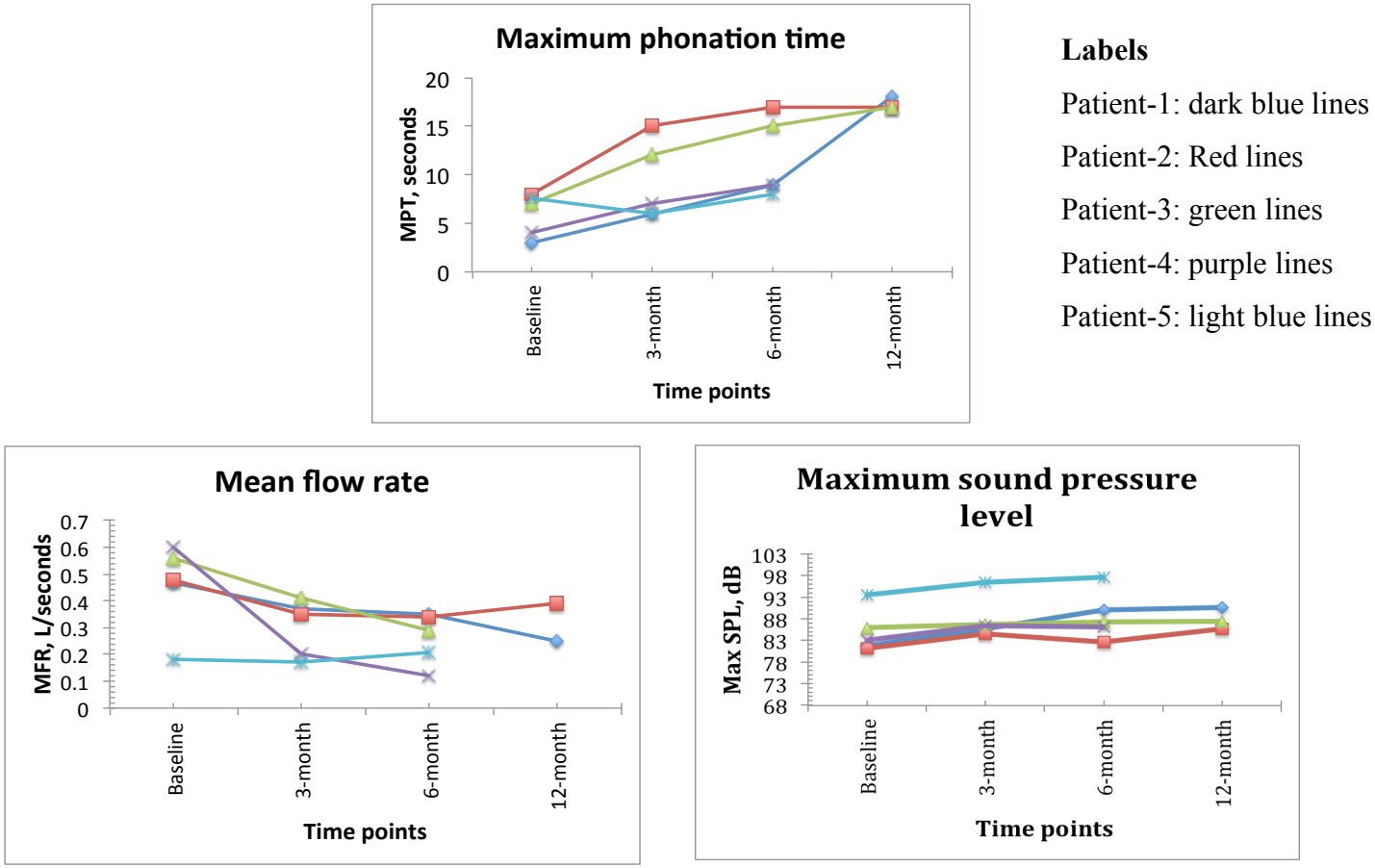


Figure 8-2: These graphs show maximum phonation time, mean flow rate and maximum sound pressure level of 5 patients.

Voice perceptual evaluation and acoustic analysis

Voice perceptual evaluation showed an improvement in the quality of voice for all patients following reinnervation. The inter- and intra-rater reliability of the voice perceptual evaluation, calculated using weighted kappa coefficient and intraclass correlation coefficient is shown in Table 8-4 and Table 8-5, and is similar to the reliability reported by Webb et al. (Webb et al., 2003). Table 8-6 showed the pre-operative and post-operative grading for 'G', 'R' and 'B' of all patients. At 12-months, the overall dysphonia of the 3 patients who completed follow-up was 0.0 (normal).

The objective voice quality evaluation; jitter percent; shimmer percent; and NHR of the patients' 5-seconds vowel/a/ at baseline, 3-, 6-month was 2.28 (1.84)%, 1.52 (1.92)% and 1.06 (0.86)% respectively for jitter; 4.95 (1.90)%, 4.27 (2.44)% and 4.47 (2.26)% respectively for shimmer; and 0.16 (0.04), 0.17 (0.09) and 0.13 (0.02) respectively for NHR. At 12-months, the jitter percent, shimmer percent and NHR of these 3 patients were 1.00%, 0.94% and 0.38% for jitter; 2.44%, 5.38% and 2.77% for shimmer; and 0.10, 0.15 and 0.14 for NHR. The jitter and shimmer improved to within normal range at 12-months post-operatively except for the patient who had the neck abscess whose shimmer was 5.58% (just outside the normal range). Figure 8-3 shows the acoustic analysis results.

Table 8-4: Inter-rater reliability for voice perceptual evaluation as calculated using weighted kappa coefficient (kappa) and intraclass correlation (ICC)

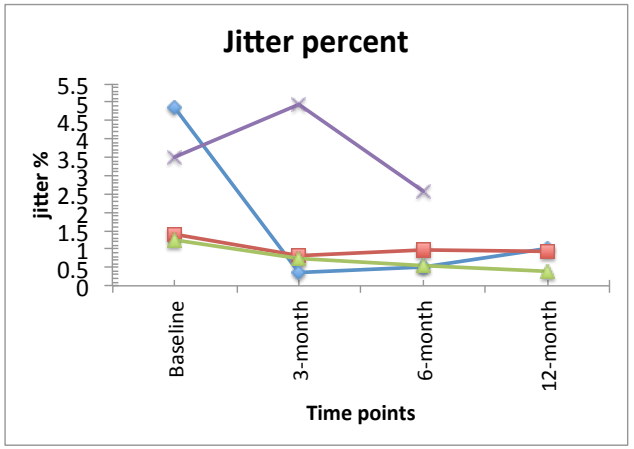
	Weighted kappa coefficient			ICC
	Rater-1 vs Rater-2	Rater-1 vs Rater-3	Rater-2 vs Rater-3	Rater-1, rater-2, rater-3
Overall dysphonia (G)	0.87	0.87	1.00	0.95
Roughness (R)	0.87	0.87	1.00	0.94
Breathiness (B)	0.95	0.95	1.00	0.98

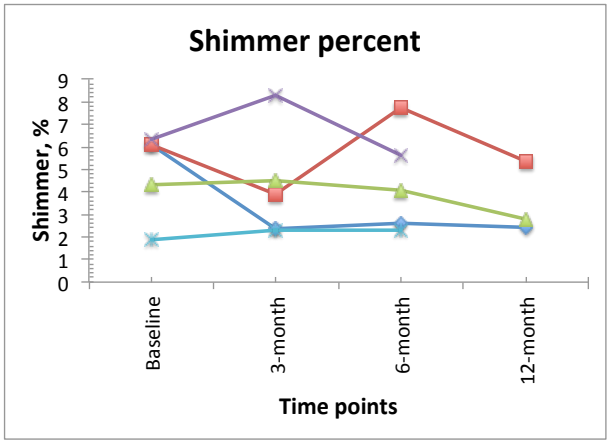
Table 8-5: Intra-rater reliability for voice perceptual evaluation as calculated using weighted kappa coefficient (kappa) and intraclass correlation (ICC)

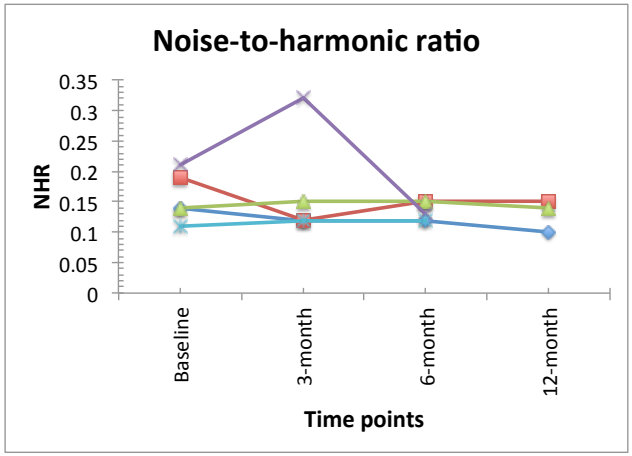
	Rater-1		Rater-2		Rater-3	
	Kappa	ICC	Kappa	ICC	Kappa	ICC
Overall dysphonia (G)	0.46	0.65	0.50	0.67	0.58	0.66
Roughness (R)	0.63	0.70	0.75	0.83	0.67	0.76
Breathiness (B)	0.39	0.62	0.35	0.59	0.35	0.59

Table 8-6: Voice perceptual evaluation by the 3 raters

	baseline	3-month	6-month	12-month
Overall dysphonia (G)				
Normal	0	3	2	3 of 3 patients
mild deviance	2	2	2	
moderate deviance	2	0	1	
Severe deviance	1	0	0	
Roughness (R)				
Normal	3	5	2	3 of 3 patients
mild deviance	1	0	3	
moderate deviance	1	0	0	
Severe deviance	0	0	0	
Breathiness (B)				
Normal	0	3	2	3 of 3 patients
mild deviance	2	2	3	
moderate deviance	2	0	0	
Severe deviance	1	0	0	







Labels

Patient-1: dark blue lines

Patient-2: Red lines

Patient-3: green lines

Patient-4: purple lines

Patient-5: light blue lines

Figure 8-3: These graphs show acoustic analysis results of 5 patients at different time points

Video-laryngostroboscopy analysis

Analysis of video-laryngostroboscopy findings was shown in Table 8-7. At 6-months, mucosal wave symmetry and duration of closure were normal in 4 patients and there was no vocal fold bowing in all patients. All 3 patients who had completed the 12-month review were normal.

Table 8-7: Video-laryngostroboscopy analysis evaluation by 3 raters

	baseline	3-month	6-month	12-month
Mucosal wave asymmetry				
No asymmetry	2	5	4	3 of 3 patients
Mildly asymmetry	1	0	1	
Moderately asymmetry	1	0	0	
Severely asymmetry	1	0	0	
Duration of closure				
Predominately closed	3	4	5	3 of 3 patients
Half-opened and half-closed	0	1	0	
Predominately open	1	0	0	
Always open	1	0	0	
Vocal fold bowing				
No bowing	3	5	4	3 of 3 patients
Mild bowing	1	0	1	
Moderate bowing	1	0	0	
Severe bowing	0	0	0	

8.1.5 Discussion

Overview

Non-selective reinnervation is one of the surgical options available to treat UVFP. The operation aims to reinnervate the paralysed vocal fold muscles and thereby restore the tone and bulk required for a good voice. However, the technique results in a delay before optimum voice improvement that is much longer than with other medialisation techniques (Aynehchi et al., 2010). This drawback may be ameliorated by temporary injection laryngoplasty, although in studies, this introduces a significant confounding factor in determining the true treatment effect of reinnervation (Lorenz et al., 2008).

Temporary injection materials for injection laryngoplasty include bovine gelatin, collagen-based products, hyaluronic acid and carboxymethylcellulose. The clinical effectiveness of these injection materials varies from 4 to 6 months (Mallur and Rosen, 2010). Wen et al. reported that porcine collagen and hyaluronic acid showed comparable results up to 6 months post-treatment during which time VHI-10 and MPT improved significantly in their retrospective case series (Wen et al., 2013). "Symptom-free duration of action" was 10.9 months for porcine collagen and 14.4 months for hyaluronic acid (Wen et al., 2013). However, the mean VHI-10 at 6 months of both groups was still abnormal (more than 11) and there were no well-validated outcome measures beyond 6 months post-treatment. The interpretation of results of this study may be compromised by biases introduced by the nature of retrospective chart review and high dropout rate of more than 20%. Those patients who lost to follow up were assumed to still having effective control of their presenting symptoms.

Moonis et al. studied longitudinal imaging appearance on T1- and T2-weighted imaging of 6 patients following injection laryngoplasty using a human based collagen. There was hyperintensity on the T1- and T2-weighted images of the injected paralysed vocal fold. The hyperintensity on T2 was more likely due to increased in fluid, protein and vasculature caused by the injection. The study suggested that injected human based collagen product persists in the vocal folds up to 11 months. However, the signal changes (hyperintensity) was not measured objectively, each radiologic assessment was not matched with the subjective and objective voice quality evaluation, and of the 6 patients, only one patient had a serial scan over time. The rest of MRI scans was done on different patients at different time points (Moonis et al., 2005).

Based on this data, I hypothesised that any voice improvements beyond 6 months following non-selective laryngeal reinnervation with concomitant injection laryngoplasty are likely to be due to the effects of reinnervation rather than of injection laryngoplasty. Here, I used LEMG as a semi-quantitative outcome measure, but also used 3T MRI as an exploratory outcome measure to assess muscles' signal changes.

Synopsis of key findings

In the present prospective case series, reinnervation improved physical, functional and emotional concerns due to UVFP compared to baseline, and this was matched by LEMG and T2-MRI outcomes, including at a 12-month time-point in 2 patients. Improvement in MPT, voice perceptual evaluation, acoustic analysis, video-laryngostroboscopy analysis and reduction of MFR were all also observed. However

patient-5 had minimal MPT improvement that may be due to limited pulmonary reserve caused by a pre-existing phrenic nerve paralysis.

In patient-2 the voice improvement was not optimum, the motor unit recruitment on LEMG was only slightly better than the baseline and the signal changes on the T2-MRI was still high at 12-month review. The patient was still unsatisfied with voice projection especially in places with high background noise. This unsatisfactory result may be due to the neck abscess post-operatively which may in turn have damaged the nascent nerve repair site. Despite this, however, at 12-months his MPT was within the normal range and voice and video-laryngostroboscopy outcomes were normal, although the voice amplitude perturbation (shimmer) remained outside the normal range.

Some of the patients in this case series were noted to have normal voice perceptual evaluation before the reinnervation even though their self-rated VHI-10 was > 20. This is because they did not have problems with individual conversation in a quiet room, but had problems in crowded places or in groups where projection was required. In these patients, at 6- and 12-month review, the VHI-10 scores reduced to normal level and the voice perceptual evaluation remained normal.

Comparison with previous works

Offering temporary voice strength by injection laryngoplasty is necessary while waiting for the reinnervation to become effective at between 3 and 4 months. Previous retrospective studies reported results of non-selective reinnervation using ansa-RLN with concomitant injection laryngoplasty using various materials-

Gelfoam (Pzifer, Cambridge, MA), micronized AlloDerm (Cymetra; LifeCell Corp, Branchburg, NJ), or microfibrillar collagen (INSTAT Microfibrillar Collagen Hemostat; Ethicon, Johnson & Johnson, Somerville, NJ). The studies reported improvement in patient self-reported voice perception, voice perceptual evaluation by professionals, acoustic analysis and video-laryngostroboscopy findings more than 6 months after reinnervation (Lee et al., 2007; Lorenz et al., 2008; Olson et al., 1998). A significant improvement in: acoustic analysis ($p<0.05$) (Olson et al., 1998), voice perceptual evaluation ($p<0.05$) (Olson et al., 1998), $p<0.001$ (Lorenz et al., 2008), and glottic closure and vocal fold bowing on video-laryngostroboscopy ($p<0.05$) (Lorenz et al., 2008) were reported. However none of these studies included LEMG results to directly assess the effect of reinnervation on laryngeal muscles. Further retrospective studies did not perform concomitant injections and observed voice outcome improvement in parallel with encouraging LEMG findings (Maronian et al., 2003; Wang et al., 2011).

The present study employing an ansa-RLN technique with concomitant collagen injection laryngoplasty is the first study to my knowledge that includes T2-MRI and correlates it with LEMG findings to assess both resorption of injection material and the return of innervation to the TA muscle and also the prospective case series to use multidimensional assessment

Limitation of the present study

The main limitation of the present study is the small number of patients and the absence of a control group. Further data collection and follow-up will be continued and updated results reported in due course.

Clinical impact

The present study provides further supporting data for the use of VHI-10 as the primary outcome measure for voice surgery trials for UVFP. T2-MRI may be a promising outcome measure for laryngeal reinnervation trials, as it may confirm denervation and reinnervation status as well as monitor resorption of injection materials.

8.1.6 Conclusion

In this prospective case series of 5 patients undergoing non-selective laryngeal reinnervation using ansa-RLN technique with concomitant injection laryngoplasty, voice improvement was demonstrated by VHI-10 and other multidimensional outcome measures, and these were correlated with positive findings on LEMG and T2-MRI.

8.2 Immediate selective laryngeal reinnervation in vagal paraganglioma patients

8.2.1 Abstract

Objective: This prospective study is to present the outcomes of immediate selective laryngeal reinnervation in improving voice, alleviating aspiration and re-establishing vocal fold mobility in patients with unilateral vagal paralysis following vagal nerve tumour excision.

Methods: Two women (40 and 52 years old) diagnosed with left cervical vagal paraganglioma undergoing excision operation were included in the study. Patient-1 have already had left vocal fold paralysis whereas patient-2 had normal vocal folds at presentation. They underwent immediate selective laryngeal reinnervation using split phrenic nerve to the abductor branch and ansa cervicalis to the adductor branch of recurrent laryngeal nerve with concomitant collagen injection laryngoplasty following vagal tumour excision. All of them were subjected to multidimensional outcome measures (voice handicap index-10: VHI-10, EAT-10, voice perceptual evaluation, acoustic analysis, maximum phonation time, aerodynamic analysis and video-laryngostroboscopy analysis) at pre-operative, 1-, 6-, and 12-months post-operative. Laryngeal electromyography (LEMG) and T2-weighted MRI (T2-MRI) of the larynx was performed at baseline and at 12-months to measure the neuromuscular integrity of the thyroarytenoid muscle and posterior cricoarytenoid muscle. Patient-1 and patient-2 baseline was before and after respectively, vagal tumour excision. Their respiratory function was assessed using MRC breathlessness scale and lung function test.

Results: Patient-1 and patient-2 had improved VHI-10 scores from of 23 and 18 respectively at 1-month, to of 5 and 1 respectively at 12-months. Their EAT-10 score was improved from of 20 and 24 respectively at 1-month to of 3 and 1 respectively at

12-months. The motor units recruitment on LEMG was ‘very much better’ for patient-2 when compared to the baseline but for patient-1, the baseline was slightly better than at-12months. There was higher signal intensity changes detected on the paralysed TA compared to the opposite ‘control’ on T2-MRI images at baseline. Repeat T2-MRI at 12-month showed normalisation of the signal intensity. There was slight vocal fold abduction for patient-1 observed and no obvious abduction for patient-2. The MRC breathlessness scale of both patients had been the same at all time points. There was no significant decreased of post-operative FEV1 or FVC from the pre-selective reinnervation measurements.

Conclusion: Selective reinnervation is a safe operation to be performed in an immediate first operated side vagal paraganglioma excision. Voice and swallowing improvement were demonstrated but no significant vocal fold abduction achieved. Reviewing patients at 1-month post-tumour excision may be the best time for baseline measurements in such studies.

8.2.2 Background

Patients with cervical vagal tumour commonly present with a neck mass and the majority of them are asymptomatic as tumour growth is gradual. The definitive treatment for this condition is surgery in which post-operative unilateral vagal paralysis ensues (Lozano et al., 2008). Unilateral vagal paralysis patients suffer dysphonia, dysphagia and aspiration. Additional operative complications of glossopharyngeal and hypoglossal nerve paralysis worsen the swallowing problems (Gilmer-Hill and Kline, 2000; Miller et al., 2000).

Multiple paraganglioma is common especially in patients with a positive family history (Urquhart et al., 1994). Netterville et al. reported that 37% of vagal paragangliomas were bilateral (Fang et al., 2011; Netterville JL et al., 1998). Multicentricity of vagal paragangliomas is a significant factor to consider in decision making of the therapeutic options (Lozano et al., 2008). Three of 19 patients with paraganglioma as reported by Urquhart et al. were subjected to irradiation therapy rather than surgery due to multicentricity. This was to avoid the incapacitating morbidity of bilateral vocal fold paralysis (BVFP) that might end up with a tracheostomy.

Significant issues of dysphonia, dysphagia and aspiration caused by unilateral vagal paralysis following vagal paraganglioma excision definitely require rehabilitation. Surgical interventions to improve voice and swallowing in this condition include thyroplasty, arytenoid adduction, injection laryngoplasty and laryngeal reinnervation (Bielamowicz et al., 2000; Fang et al., 2011; Lamarre et al., 2011; Miller et al., 2000; Urquhart et al., 1994). These studies had shown the usefulness of the surgical

interventions in rehabilitating the voice and swallowing issues. However their results were limited by the retrospectivity of the studies, inherent biases and the lack of details of voice and swallowing treatment outcome monitoring effect.

Surgical interventions are ideally performed at the same sitting to minimise post-operative morbidity. Injection laryngoplasty using current available substances usually provides only a temporary effect, necessitating repeated injections. Thyroplasty is commonly done under local anaesthesia since patients' cooperation is necessary to fine-tune the voice according to implant size and placement. Laryngeal reinnervation may be a good option since the operation is preferably done under general anaesthesia and it does not require fine-tuning of voice. The surgical reinnervation that aims to re-establish tone and bulk of the denervated muscle may be better than other surgical techniques as Woodson in 2007 had shown in cats that spontaneous regeneration did not occur in vagus transection possibly due long course of the nerve to RLN and only small proportion of the axons goes to RLN (Woodson, 2007). Further more, re-establishment of vocal fold mobility may be possible by performing selective reinnervation that may save patients from irradiation therapy or tracheostomy should the tumour grow later on the opposite neck. A prospective study with standardised outcome measures is necessary to demonstrate the effect of selective reinnervation in unilateral vagal paralysis following the first operated side of vagal paraganglioma.

This prospective study is to present the outcomes of selective laryngeal reinnervation in improving voice, alleviating aspiration and re-establishing vocal fold mobility in patients with unilateral vagal paralysis following vagal nerve tumour excision.

8.2.3 Methods

Patient selection

Two female patients diagnosed with left cervical vagal paraganglioma from ENT clinics of Charing Cross Hospital undergoing excision operation were included in the study. Patient-1 (40 years old) have already had left vocal fold paralysis at presentation to the clinic whereas patient-2 (52 years old) had normal vocal folds at presentation. Neither of them had any other medical problems.

Outcome measures

The patients were assessed pre- and post-operatively for voice improvement using multidimensional measures, effect of reinnervation on the thyroarytenoid muscles, swallowing and respiratory condition as the following:

a) voice handicap index-10 (VHI-10);
b) voice perceptual evaluation (GRBAS);
c) acoustic analysis (jitter(%), shimmer(%), NHR);
d) aerodynamic analysis (maximum SPL(dB), mean SPL(dB), mean airflow rate at voicing(MFR, L/s));
e) video-laryngostroboscopy (mucosal asymmetry, duration of closure, vocal fold bowing; vocal fold abduction);
f) swallowing questionnaire (EAT-10);
g) Medical Research Council (MRC) breathlessness scale;
h) pulmonary function test (FEV1, FVC);
i) laryngeal electromyography (LEMG); and

j) thyroarytenoid muscles (TA) signal intensity and posterior cricoarytenoid (PCA bulk) on T2 weighted MRI of larynx (T2-MRI).

Outcome measures for no. 1 to 8 above were done at pre-operative and at 1, 6 and 12 months post-operative whereas the LEMG and T2-MRI for neurological status assessment were done at baseline, 6 and 12 months after operation. For the LEMG and T2-MRI, the baseline point was different for patient-1 and patient-2. The patient-1 had the baseline tests done before the surgery (tumour excision plus selective reinnervation plus injection laryngoplasty that was done at the same sitting). This is because patient-1 had already had UVFP before the tumour excision. Whereas, patient-2 who had pre-operative normal vocal folds, had the LEMG and T2-MRI done 5 weeks after the surgery (tumour excision plus injection laryngoplasty) that was 1 week before the selective reinnervation. The selective reinnervation that was initially planned for the same sitting as that for tumour excision for patient-2 had to be done later due to unavoidable circumstances.

For details of description of outcome measures, please refer chapter 8.1, page 247 to 249.

Selective reinnervation – surgical technique

Selective laryngeal reinnervation was done following a vagal nerve tumour operation either at the same setting or 5 weeks following the excision operation. A separate incision was made on the neck lower to the first incision for vagal nerve tumour removal. The separate incision followed the skin crease. Thyropharyngeus muscle

overlying the lateral border of thyroid cartilage lamina was divided for the purpose of intra-laryngeal RLN dissection. The main trunk of the RLN was identified behind the cricothyroid joint. The abductor and adductor branch of the RLN were then identified. Phrenic nerve was normally found deep to the transverse cervical artery overlying the anterior scalenus muscle. Split phrenic nerve technique was used to reinnervate the PCA muscle. Greater auricular nerve was used as a cable graft to bridge the phrenic nerve and the distal stump of the RLN before it branched into the abductor branch. The adductor branch of RLN was anastomosed to the ipsilateral ansa cervicalis nerve (Crumley and Izdebski, 1986). Injection to the left vocal fold using porcine collagen (Permacol™, Tissue Science Laboratories plc, "TSL" Aldershot UK) was done to temporarily help the voice while waiting for the reinnervation to come into effect.

8.2.4 Results

Results of the two patients' vocal assessments, swallowing assessments and aerodynamic analysis at before operation, 1-, 6- and 12-months after operation were summarised in Table 8-8. For neurological and pulmonary function status assessments, the data was summarised in Table 8-9and Table 8-10 respectively.

Patient-1

This patient complained of hoarseness for 12 months before the diagnosis of left cervical vagal paraganglioma. Her VHI-10 was 16/40 and her voice was grade 2 for overall dysphonia and MPT was 15s. For aerodynamic analysis, MFR at voicing was 0.11L/s and maximum SPL was 90.81dB. On video-laryngostroboscopic examination, the left vocal fold was immobile, in paramedian position although no mucosal asymmetry, normal duration of vocal fold closure and no significant bowing of the vocal fold edge noted (Figure 8-4). She had slight dysphagia and aspiration that her EAT-10 was 6/40. LEMG of TA showed small amplitude motor unit that firing at high rate in isolation that may represent conduction block. T2-MRI showed increased signal intensity on the left TA (Figure 8-7). She did not complain any significant difficulty in breathing (MRC breathlessness scale, grade 2) that she had breathlessness only when she was hurrying on the level or climbing up a slight hill. Her chest X-ray findings and pulmonary function test results were normal.

She had selective reinnervation and collagen injection at the same sitting with the excision of vagal paraganglioma. There was evidence of left diaphragmatic paralysis post-operatively on the chest X-ray. She was on nasogastric tube feeding for about a week and then discharged home with oral feeding. The tumour excision was

complicated by left side tongue and soft palate weakness. At 1-month review, her voice was grade 1 for overall dysphonia and the VHI-10 was 23/40. Her EAT-10 was 20/40 and she managed semi-solid food. Her MRC breathlessness scale did not change from the pre-operative state.

At 6 and 12 months review, her swallowing continued to improve after the selective reinnervation. Her voice was mildly dysphonic (G, 1). LEMG showed polyphasic pattern (large) that was consistent with reinnervation. However, the motor unit recruitment at baseline was slightly better than the post-operative. T2-MRI TA signal intensity and PCA bulk of the paralysed side reduced and increased respectively to almost similar to the opposite normal side (Figure 8-7, Figure 8-8 and Figure 8-9). Her VHI-10 and EAT-10 were within normal limits (Figure 8-6). The overall grade for dysphonia was grade 1 and her MPT was 10s. Her pulmonary symptoms and function remain unchanged.

Acoustic analysis results and maximum SPL had been within normal limits throughout all time points. The MFR was worst at 6-months (0.31 L/s) and improved to 0.21 L/s at 12-months. There was slight left vocal fold abduction re-established.

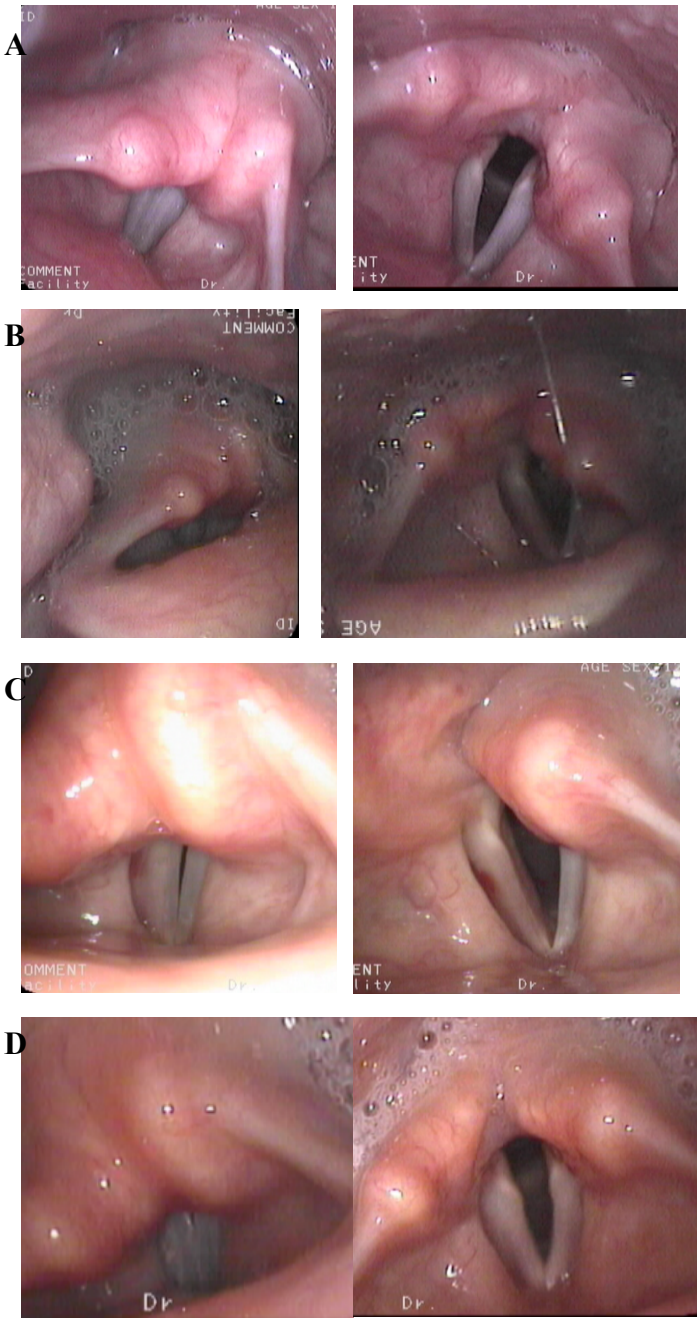


Figure 8-4: Videostroboscopy images of patient-1 during phonation (left column) and breathing (right column) at: (A) before; (B) 1-; (C) 6-; and (D) 12-months post selective reinnervation

Patient-2

This patient complained of left neck mass and was diagnosed left vagal paraganglioma. She did not have any voice or swallowing issues by which all of her clinical assessment parameters were within normal limit. She was planned for selective reinnervation at the same sitting as that for the vagal paraganglioma excision. However due to unavoidable circumstances, only vocal fold injection with collagen was performed. She was on nasogastric tube feeding for about a week after the tumour excision and then discharged home with oral feeding.

Selective reinnervation was done 5 weeks after the vagal paraganglioma excision. At 4-weeks post-excision and collagen injection, before the selective reinnervation, her VHI-10 was 18/40, overall grade for dysphonia was grade 1, EAT-10 was 24/40, and MPT was 10s. LEMG of the TA revealed absent motor unit recruitment with fibrillation potential. T2-MRI showed increased signal intensity of the left TA (paralysed) but there was no significant reduction in the left PCA bulk (Figure 8-7 and Figure 8-9). Her chest X-ray findings and pulmonary function test results were normal. She did not complain any significant difficulty in breathing (MRC breathlessness scale, grade 2) in which she had breathlessness when she was hurrying on the level or climbing up a slight hill. Immediate post-selective reinnervation, there was evidence of left diaphragmatic paralysis post-operatively on the chest X-ray.

At 6- and 12-months, her voice and swallowing continue to improve (Figure 8-6). LEMG showed polyphasic pattern that consistent with reinnervation and the recruitment was very much better than the baseline. T2-MRI TA signal intensity and PCA bulk of the paralysed side reduced and increased respectively to almost similar

to the opposite normal side (Figure 8-7, Figure 8-8 and Figure 8-9). At 12-months, her VHI-10 and EAT-10 were within normal limit. The overall grade for dysphonia was grade 0 and her MPT was 10s. Her pulmonary symptoms and function remain unchanged.

Acoustic analysis results and maximum SPL had been within normal limit throughout all time points. The MFR was worst at 1-month (0.37 L/s) and improved to 0.27L/s at 12-months. There was no significant left vocal fold movement re-established.

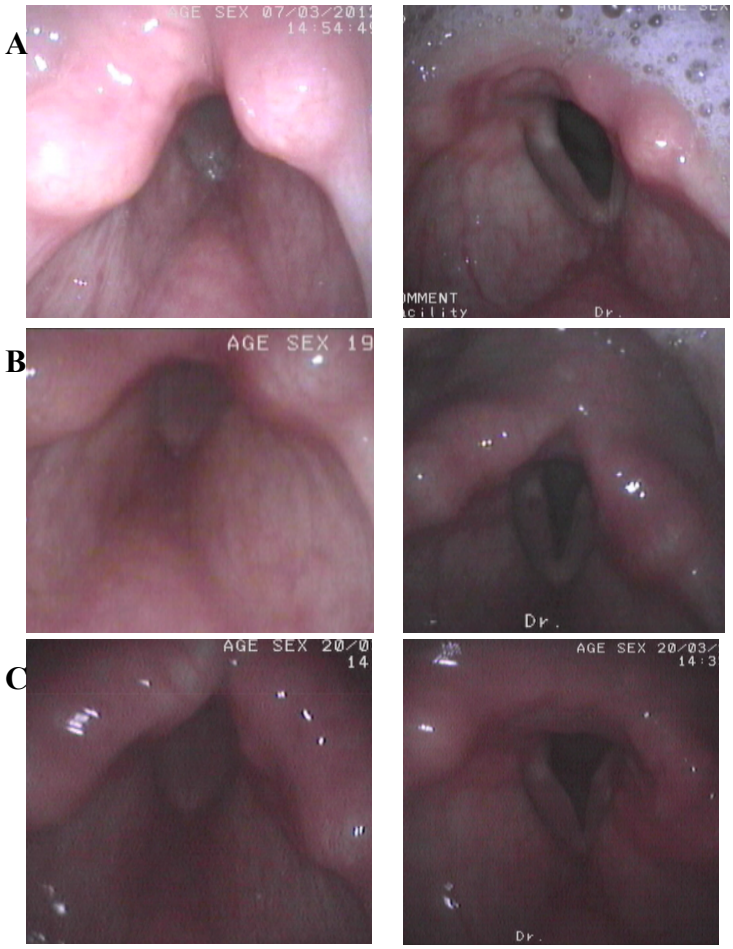
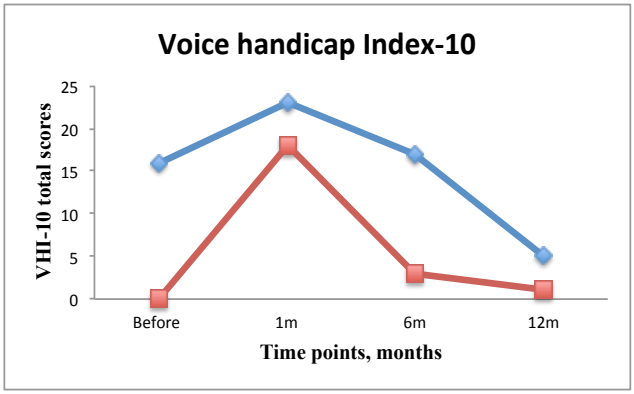


Figure 8-5: Videostroboscopy images of patient-2 during phonation (left column) and breathing (right column) at: (A)1-; (B) 6-; and (C) 12-months post-selective reinnervation. Her normal vocal folds before operation is not shown here.



Labels

Patient-1: blue lines

Patient-2: red lines

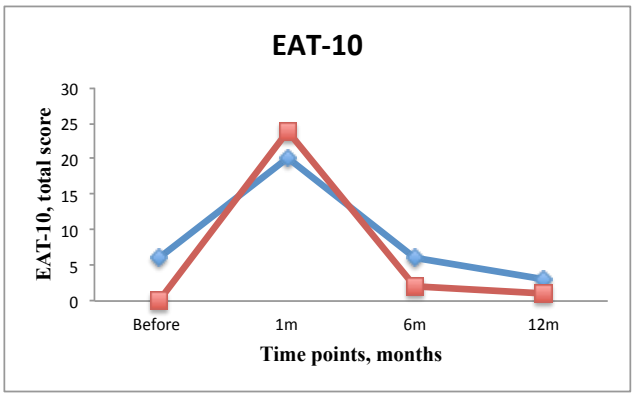


Figure 8-6: Graphs depict improvement of voice and swallowing in patient-1 and patient-2 as measured by the VHI-10 and EAT-10.

Table 8-8: Summary of voice assessments, acoustic analysis and aerodynamic analysis of patient-1 and patient-2.

	Patient-1				Patient-2			
	Pre-operative	1m	6m	12m	Pre-operative	1m	6m	12m
VHI-10	16	23	17	5	0	18	3	1
Overall dysphonia grade (G)	2	1	2	1	0	1	0	0
Roughness (R)	2	1	1	0	0	1	0	0
Breathiness (B)	1	1	2	1	0	1	0	0
Jitter	1.46	1.44	2.41	1.4	0.53	1.09	0.86	0.54
Shimmer	2.71	1.28	1.9	1.67	1.76	3.54	2	2.36
NHR	0.1	0.08	0.1	0.07	0.12	0.11	0.08	0.09
MPT	15	10	9	10	15	10	9	9
max SPL	90.81	88.83	87.83	84.73	87.4	80.56	92.78	89.12
mean flow rate (L/s)	0.11	0.18	0.3	0.21	0.19	0.37	0.15	0.27
Mucosal asymmetry	0	0	0	0	0	2	0	0
Duration of closure	0	0	1	0	0	1	0	0
Vocal fold bowing	0	0	1	0	0	1	0	0
EAT-10	6	20	6	3	0	24	2	1

Table 8-9: Results of T2-MRI and laryngeal electromyography

	Patient-1				Patient-2			
	Right side		Left side (paralysed)		Right side		Left side (paralysed)	
	baseline	12m	baseline	12m	baseline	12m	baseline	12m
TA signal intensity	1.88	1.67	2.40	1.55	1.94	1.90	2.76	1.73
PCA bulk	0.31	0.33	0.27	0.32	0.32	0.30	0.21	0.25
Comparison motor unit recruitment at 12 months Between baseline and post-operative	baseline slightly better than post-operative				Post-operative very much better than baseline			

Table 8-10: Pulmonary function status at before and after selective reinnervation

	Patient-1					Patient-2			
	before	1m	6m	12m		before	1m	6m	12m
MRC scale	2	2	2	2		2	2	2	2
FEV1	2.61	Not done	2.59	2.57		2.97	Not done	2.86	2.97
FVC	3.39	Not done	3.4	3.44		4	Not done	3.91	4.00
FEV1/FVC (%)	77	Not done	76.08	74.72		74.25	Not done	73.02	74.18

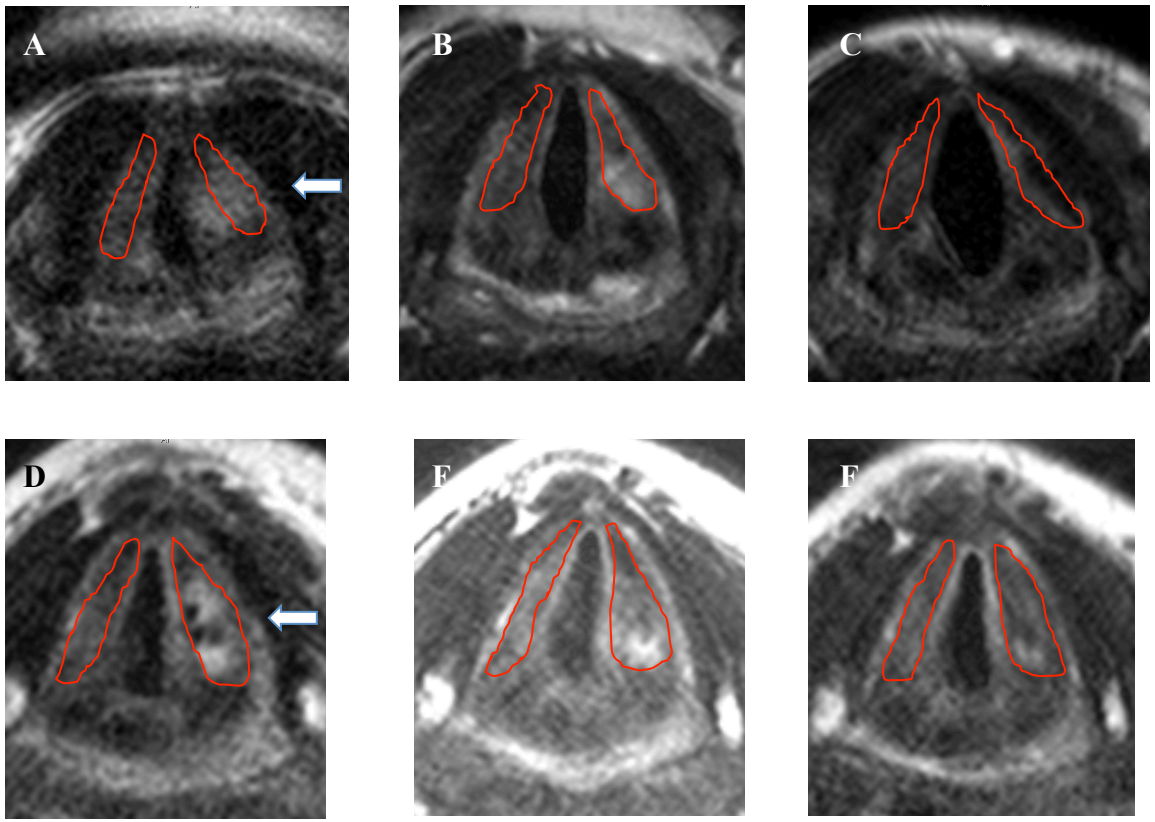


Figure 8-7: Patient-1 (top row) and patient-2 (bottom row), T2-MRI showing signal intensity changes on the left TA muscle at baseline (A,D); 6- (B,E); and 12-months (C,F) following the reinnervation. The arrow shows the left TA.

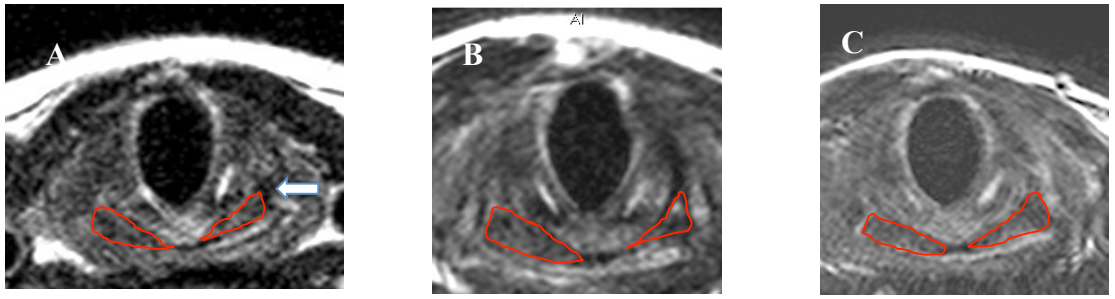


Figure 8-8: Patient-1 T2-MRI PCA bulk on axial plane at: baseline (A); 6- (B); and 12-months (C) post-selective reinnervation. The arrow shows the paralysed side.

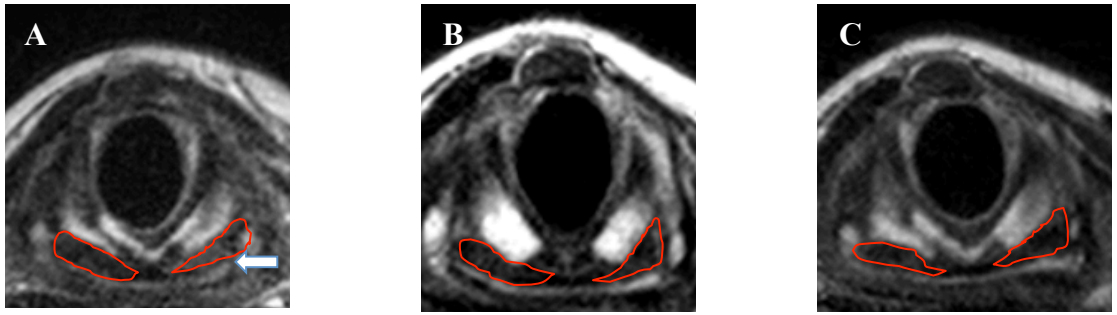


Figure 8-9: Patient-2 T2-MRI depicting PCA bulk on axial plane at baseline (A); 6- (B); and 12-months (C) post-selective reinnervation. The arrow shows the paralysed side.

8.2.5 Discussion

Overview

Vagal paragangliomas are rare but cause significant post-tumour excision morbidity as a result of unilateral vagal paralysis (Netterville JL et al., 1998). Multiple paragangliomas are common especially in those with a positive family history, and this results treatment dilemma: a later synchronous contralateral lesion will cause airway obstruction due to BVFP (Lozano et al., 2008). Selective reinnervation has been reported as a treatment option for BVFP to re-establish vocal fold abduction to improve the glottal airway as well preserving a good voice. The operation involves reinnervation of the PCA muscles with the phrenic nerve and adductor muscles with the ansa cervicalis nerve (ansa-RLN). Therefore selective laryngeal reinnervation is a potential attractive option to rehabilitate the post-operative morbidities.

Synopsis of key findings

Both patients demonstrated improvement of voice and swallowing following the selective reinnervation. The patients' perception on physical, functional and emotional effect of voice (VHI-10) and swallowing assessment (EAT-10) returned to normal range at 12-months post-reinnervation. At 1-month up to 6-months after the tumour excision, their voice and swallowing were helped temporarily by bulking up the vocal fold with the collagen injection. Patient-2 had better and earlier voice and swallowing improvement than patient-1 because the latter had left side tongue and soft palate weakness as well. The soft palate weakness causes hypernasality and may contribute to the mild abnormality of overall dysphonia and breathiness score component of voice perceptual evaluation. Both of the patients were on enteral

feeding only for a week and none of them had aspiration pneumonia throughout follow-up.

LEMG and T2-MRI findings showed improvement of neurological status that correlate with the voice and swallowing improvement. For patient-2, the TA motor unit recruitment was slightly better at baseline than 12-months because the baseline test was done prior to tumour removal. The vagus nerve was still intact but had conduction block causing vocal fold paralysis. Therefore the motor units morphology and recruitment appeared healthier than after the vagal tumour excision. The T2-MRI showed reduced TA signal intensity and increased PCA bulk at 12-months. The reversal of signal intensity may indicate reinnervation of the paralysed muscles as well as resorption of the collagen substance that was injected into the vocal fold immediately after the vagal paraganglioma excision.

For glottal airway, there was slight abduction for patient-1 observed and no obvious vocal fold abduction for patient-2. Nevertheless using phrenic nerve for selective reinnervation did not cause significant morbidity to the pulmonary function in the present study. The MRC breathlessness scale of both patients had been the same at all time points. There was no obvious decreased of post-operative FEV1 or FVC from the pre-selective reinnervation measurements.

Strength of the study

This is a prospective study of case series that includes multidimensional analysis to monitor voice and swallowing outcomes of unilateral vagal paralysis patients who had primary or immediate selective reinnervation following vagal paraganglioma

excision. The success of reinnervation in improving voice and swallowing was correlated with evidence of reinnervation on LEMG and T2-MRI findings.

Comparison with previous studies

Laryngeal reinnervation (non-selective) has been reported to improve voice and swallowing in unilateral vagal paralysis following vagal paraganglioma excision. Lee et al. and Lorenz et al. included this group of patients in their retrospective studies whereby they had ansa-RLN reinnervation immediately after the tumour removed. They documented favourable outcome of ansa-RLN anastomosis characterised by improvement in patients' voice perception, acoustic analysis, long lasting improvement of the glottic closure and maintenance of the vocal fold edge (Lee et al., 2007; Lorenz et al., 2008). However in these studies, data analysis of those patients was problematic due to unavailability of the preoperative data since the voice was normal before the tumour resection.

Lamarre et al. reported a case of vagal paraganglioma resection with primary reinnervation of the larynx in the same sitting. In this case report, apart from ansa-RLN anastomosis, cricothyroid-to-cricothyroid reinnervation and greater auricular nerve-to-superior laryngeal nerve anastomosis were performed. The patient showed good glottic closure at 12-months postoperatively and recovered swallowing function but the measurements was not explicitly reported. To date there are no published attempts at selective reinnervation to the PCA to re-establish vocal fold abduction in this group of patients.

Limitation of study

This study only included two patients and there was no control arm. The small number of subjects was due to the rare nature of the tumour that is about 9.1% of head and neck paragangliomas (Lozano et al., 2008; Netterville JL et al., 1998). However it provides comprehensive and prospective treatment monitoring effect.

Selective reinnervation in the present study successfully improved the voice and swallowing but failed to re-establish significant vocal fold abduction. This may be due laryngeal synkinesis or inability of axons from the phrenic nerve to trigger the abduction movement. Nevertheless, the paralysed PCA exhibited reinnervation changes indicated by increased in muscle bulk on T2-MRI. However there was no PCA LEMG data available to verify this due to technical challenges in assessing the muscles. Although PCA and TA are both supplied by the RLN, the degree of reinnervation may be different as the PCA is supplied by the abductor axons and the TA is supplied by the adductor axons. The presence of laryngeal synkinesis was also not included in the TA LEMG evaluation.

Clinical applicability

This study demonstrates the applicability of performing immediate selective reinnervation or at the same sitting in vagal paraganglioma excision to reduce the post-operative morbidities. It re-established the tone and bulk of the TA thus improved the voice to normal or near normal as well as swallowing. Theoretically, a successful re-establishment of vocal fold abduction will save future dilemma of treatment decision making should the tumour grow on the opposite side. Phrenic nerve split arguably should be replaced with another technique introduced by Marie

that uses root of the phrenic nerve (C4) (Marina et al., 2011; Remacle and Eckel, 2010).

8.2.6 Conclusion

Selective reinnervation (split phrenic nerve to abductor branch and ansa cervicalis to adductor branch) is a safe operation to be performed in an immediate first operated side vagal paraganglioma excision. Voice and swallowing improvement were demonstrated but no significant vocal fold abduction achieved. Future studies probably should use the root of the phrenic nerve instead of the split phrenic nerve technique. Reviewing patients at 1-month post-tumour excision may be the best time for baseline measurements in such studies.

General conclusion

The unifying theme of this thesis is a series of research studies collectively representing a feasibility study for clinical trials of laryngeal reinnervation for the treatment of vocal fold paralysis.

In chapter 2, I have shown that an RCT of reinnervation versus thyroplasty in patients with UVFP is feasible in the UK. The results of the ENT UK survey and qualitative study of patients with UVFP, also highlight potential problem areas for recruitment of surgeon-investigators and patients to this trial. A potential pool of patients for the trial appears promising. The majority of ENT surgeons in the UK who manage patients with UVFP are receptive to the proposed trial and are willing to be trained. Training on laryngeal reinnervation is vital as not many surgeons are familiar with this surgery. The data of individual interview of patients presented evaluates patient-focus and optimisation of the trial protocol, and both recruitment and consent processes. I found phraseology that needed changing or avoiding during the recruitment process. This may in turn improve willingness of potential patients to be randomised. I also proposed VHI-10 to be used as the primary outcome measure in the proposed RCT.

In chapter 3, I have shown that OperaVOX (a potential new software package that was installed in an iPod touch) is statistically comparable to the 'gold standard' (MDVP) for most principal phonatory outcome measures. However, given its portability, low cost and applicability to home or clinic, OperaVOX has greater utility and therefore may be preferred for voice outcome data collection in both

settings. Similarly, it may be preferred by patients and clinicians for routine clinical data collection.

In chapter 4, I demonstrated the reliability of video-laryngostroboscopy as a tool for evaluating UVFP. A substantial inter- and intra-rater reliability was shown for the mucosal wave asymmetry, duration of closure and vocal fold bowing of the video-laryngostroboscopy parameters that appear to be reliable parameters for clinical and research use. These parameters are important in indicating return of voice function by closing the glottal gap as well as the preservation of pliability of the vocal fold mucosa, for example that following laryngeal reinnervation.

In chapter 5, I demonstrated the repeatability and reproducibility of T2-MRI as a non-invasive tool in depicting denervation changes in vocal fold paralysis muscles. Signal changes on TA muscles had shown to be correlated with electrophysiological results on LEMG that is useful in monitoring treatment outcome of laryngeal reinnervation in UVFP patients. Other metrics may be useful as well and better reliability may be demonstrated if a bigger number of participants are included. A larger scale study that includes more volunteers and patients is necessary to decide on a normal cut off range.

In chapter 6, I examined the feasibility of assessing tissue perfusion in laryngeal muscles using DCE-MRI and DWI. The findings confirm that paralysed TA and PCA muscles have reduced perfusion and that apparently normal contralateral musculature in patients with UVFP has elevated perfusion. This work will inform mechanistic studies of vascular change in paralysed laryngeal muscles and clinical

studies evaluating utility perfusion parameters in predicting and monitoring responses to treatment.

In chapter 7, I have shown that cine-MRI quantitative metrics are feasible to objectively quantify vocal fold mobility and had high repeatability (except VFPa) for identification of patients with UVFP. I propose that this technique would be of great benefit to non-invasively and objectively assess the graded improvement in vocal fold mobility that are likely to occur following surgical interventions aimed at re-establishing the dynamic movement of vocal fold.

In chapter 8.1, I have presented a prospective case series of 5 patients undergoing non-selective laryngeal reinnervation using ansa-RLN technique with concomitant injection laryngoplasty, voice improvement was demonstrated by VHI-10 and other multidimensional outcome measures, and these were correlated with positive findings on LEMG and T2-MRI. The present study provides further supporting data for the use of VHI-10 as the primary outcome measure for voice surgery trials for UVFP. T2-MRI may be a promising outcome measure for laryngeal reinnervation trials, as it may confirm denervation and reinnervation status as well as monitor resorption of injection materials.

In chapter 8.2, I have reported that selective reinnervation (split phrenic nerve to abductor branch and ansa cervicalis to adductor branch) is a safe operation to be performed in an immediate (first operated side) vagal paraganglioma excision. Voice and swallowing improvement were demonstrated but no significant vocal fold abduction achieved. Future studies probably should use the root of the phrenic nerve

instead of the split phrenic nerve technique. Reviewing patients at 1-month post-tumour excision may be the best time for baseline measurements in such studies.

In summary, I have shown that a randomised trial of laryngeal reinnervation versus thyroplasty is feasible in the UK, and have validated patient- and observer-rated outcome measures. I have also shown that MRI may offer an alternative to electromyography in the assessment of laryngeal neuromuscular function in future trials and the clinic.

General limitations and future research

The ENT UK survey results were promising but the surgeons' equipoise in management of patients with UVFP is unknown. Another survey that employs surgeons' equipoise measurement is necessary to address ENT surgeons, members of British Laryngological Association who have main interest in managing patients with UVFP.

The qualitative study interviewing eligible patients for the proposed RCT gave some input on the percentage of willingness for randomisation. However, the results cannot be generalised as the recruitment centre was only in a few hospitals in London. A multi-centre pilot study of the RCT with integrated qualitative research is necessary to confirm the feasibility also the sensitivity of the proposed primary outcome measure (VHI-10).

T2-MRI images are susceptible to motion artefacts. Participants were not allowed to clear their throat or swallow during acquisition period to ensure good quality of images. Some UVFP patients may find it difficult especially when they have pooling of saliva due to swallowing issues. This issue may be overcome by improving the time acquisition in the protocol.

The T2-MRI study had included small number of patients with a limited range of neurological status on LEMG results. Therefore, based on this study, I could not make a conclusion whether T2-MRI anatomical metrics can be used to demonstrate the severity of paralysed muscles' neurological status. Future studies should recruit

more number of patients with wider temporal range, from acute phase to chronic phase of denervation.

In the DCE-MRI and DWI study, as commensurate with the nature of the patients included within this study I did not obtain histological correlates of MRI findings. Further pre-clinical work in this area remains to be performed.

References

Abraham, N.S., Young, J.M., and Solomon, M.J. (2006). A systematic review of reasons for nonentry of eligible patients into surgical randomized controlled trials. Surgery *139*, 469–483.

Ahmad, M., Dargaud, J., Morin, A., and Cotton, F. (2006). Functional morphology of phonation evaluated by dynamic MRI. Surgical and Radiologic Anatomy *28*, 481–485.

Ahmad, M., Dargaud, J., Morin, A., and Cotton, F. (2009). Dynamic MRI of larynx and vocal fold vibrations in normal phonation. J Voice *23*, 235–239.

Alderson, P. (1996). Equipoise as a means of managing uncertainty: personal, communal and proxy. J Med Ethics *22*, 135–139.

Amin, M.R., Lazarus, C.L., Pai, V.M., Mulholland, T.P., Shepard, T., Branski, R.C., and Wang, E.Y. (2012). 3 Tesla turbo-FLASH magnetic resonance imaging of deglutition. Laryngoscope *122*, 860–864.

Amir, O., Wolf, M., and Amir, N. (2009). A clinical comparison between two acoustic analysis softwares: MDVP and Praat. Biomedical Signal Processing and Control *4*, 202–205.

Anderson, T.D., and Sataloff, R.T. (2004). Complications of collagen injection of the vocal fold: report of several unusual cases and review of the literature. J Voice *18*, 392–397.

Anderson, T.D., Spiegel, J.R., and Sataloff, R.T. (2003). Thyroplasty revisions: frequency and predictive factors. Journal of Voice *17*, 442–448.

Applebaum, E.L., Allen, G.W., and Sisson, G.A. (1979). Human Laryngeal Reinnervation. The Laryngoscope *89*, 1784???1787.

Arffa, R.E., Krishna, P., Gartner-Schmidt, J., and Rosen, C.A. (2012). Normative Values for the Voice Handicap Index-10. Journal of Voice *26*, 462–465.

Aynehchi, B.B., McCoul, E.D., and Sundaram, K. (2010). Systematic review of laryngeal reinnervation techniques. Otolaryngol Head Neck Surg *143*, 749–759.

Baken, R. (1996). Dynamical disorders of voice: a chaotic perspective on vocal irregularities. In Diagnosis and Treatment of Voice Disorders, Plural Publishing, San Diego, pp. 125–141.

Baldissera, F., Cantarella, G., Marini, G., Ottaviani, F., and Tredici, G. (1989). Recovery of Inspiratory Abduction of the Paralyzed Vocal Cords After Bilateral Reinnervation of the Cricoarytenoid Muscles by One Single Branch of the Phrenic Nerve. The Laryngoscope *99*, 1286???1292.

Baylor, C., Yorkston, K., Strand, E., Eadie, T., and Joseph, D. (2005). Measurement of treatment outcome in unilateral vocal fold paralysis: a systematic review (Seattle:Academy of Neurologic Communication Disorders and Sciences (ANCDS) and American Speech-Language Hearing Association (ASHA)).

Bendszus, M., Koltzenburg, M., Wessig, C., and Solymosi, L. (2002). Sequential MR Imaging of Denervated Muscle: Experimental Study. AJNR Am J Neuroradiol *23*, 1427–1431.

Bendszus, M., Wessig, C., Solymosi, L., Reiners, K., and Koltzenburg, M. (2004). MRI of peripheral nerve degeneration and regeneration: correlation with electrophysiology and histology. Exp. Neurol. *188*, 171–177.

Benninger, M.S., and Hseu, A. (2011). Laser surgical management of bilateral vocal fold immobility. Operative Techniques in Otolaryngology-Head and Neck Surgery *22*, 116–121.

Benninger, M.S., Crumley, R.L., Ford, C.N., Gould, W.J., Hanson, D.G., Ossoff, R.H., and Sataloff, R.T. (1994). Evaluation and treatment of the unilateral paralyzed vocal fold. Otolaryngol Head Neck Surg *111*, 497–508.

Benninger, M.S., Gillen, J.B., and Altaian, J.S. (1998). Changing etiology of vocal fold immobility. The Laryngoscope *108*, 1346–1350.

Bentsianov, B.L., Boruk, M., and Rosenfeld, R.M. (2002). Evidence-based medicine in otolaryngology journals. Otolaryngology - Head and Neck Surgery *126*, 371–376.

Bevan, E., Chee, L., McGhee, S., and McInnes, G. (1993). Patients' attitudes to participation in clinical trials. British Journal of Clinical Pharmacology *35*, 204–207.

Biazar, E.,Korasani, M.T.,Montazeri, N., Pourshamsian, K., Daliri, M., Rezaei, T.M., Jabarvand, B.M, Khoshzaban, A., Heidari, K.S., Jafarpour, M., Roviemiab, Z. (2010). Types of neural guides and using nanotechnology for peripheral nerve reconstruction. The International Journal of Nanomedicine *5*, 839-852.

Bielamowicz, S., and Stager, S.V. (2006). Diagnosis of unilateral recurrent laryngeal nerve paralysis: laryngeal electromyography, subjective rating scales, acoustic and aerodynamic measures. Laryngoscope *116*, 359–364.

Bielamowicz, S., Kreiman, J., Gerratt, B.R., Dauer, M.S., and Berke, G.S. (1996). Comparison of Voice Analysis Systems for Perturbation Measurement. J Speech Hear Res *39*, 126–134.

Bielamowicz, S., Gupta, A., and Sekhar, L.N. (2000). Early arytenoid adduction for vagal paralysis after skull base surgery. Laryngoscope *110*, 346–351.

Le Bihan, D., Turner, R., Moonen, C.T.W., and Pekar, J. (1991). Imaging of diffusion and microcirculation with gradient sensitization: Design, strategy, and significance. Journal of Magnetic Resonance Imaging *1*, 7–28.

Bland, J.M., and Altman, D.G. (1999). Measuring agreement in method comparison studies. Stat Methods Med Res *8*, 135–160.

Blitzer, A., Crumley, R.L., Dailey, S.H., Ford, C.N., Floeter, M.K., Hillel, A.D., Hoffmann, H.T., Ludlow, C.L., Merati, A., Munin, M.C., et al. (2009).

Recommendations of the Neurolaryngology Study Group on laryngeal electromyography. Otolaryngology - Head and Neck Surgery *140*, 782–793.

Blumin, J.H., and Merati, A.L. (2008). Laryngeal reinnervation with nerve-nerve anastomosis versus laryngeal framework surgery alone: A comparison of safety. Otolaryngology - Head and Neck Surgery *138*, 217–220.

Borges, A. (2010). Imaging of denervation in the head and neck. European Journal of Radiology *74*, 378–390.

Borisov, A.B., Huang, S.-K., and Carlson, B.M. (2000). Remodeling of the vascular bed and progressive loss of capillaries in denervated skeletal muscle. The Anatomical Record *258*, 292–304.

Bough Jr., I.D., Heuer, R.J., Sataloff, R.T., Hills, J.R., and Cater, J.R. (1996). Intrasubject variability of objective voice measures. Journal of Voice *10*, 166–174.

Brancatisano, T., Collett, P.W., and Engel, L.A. (1983). Respiratory movements of the vocal cords. J Appl Physiol *54*, 1269–1276.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology *3*, 77–101.

Breyer, T., Echternach, M., Arndt, S., Richter, B., Speck, O., Schumacher, M., and Markl, M. (2009). Dynamic magnetic resonance imaging of swallowing and laryngeal motion using parallel imaging at 3 T. Magn Reson Imaging *27*, 48–54.

Brockmann-Bauser, M., and Drinnan, M.J. (2011). Routine acoustic voice analysis: time to think again? Curr Opin Otolaryngol Head Neck Surg *19*, 165–170.

Calvert, M., Blazeby, J., Revicki, D., Moher, D., and Brundage, M. (2011). Reporting quality of life in clinical trials: a CONSORT extension. Lancet *378*, 1684–1685.

Carding, P., Carlson, E., and Eps, R. (2000). Formal perceptual evaluation of voice quality in the United Kingdom. Logoped Phoniatr Vocol *25*, 133–138.

Carding, P., Steen, I. N., Webb, A., MacKenzie, K., Deary, I. J., and Wilson, J. A. (2004). The reliability and sensitivity to change of acoustic measures of voice quality. Clinical Otolaryngology & Allied Sciences *29*, 538–544.

Carding, P.N., Wilson, J.A., MacKenzie, K., and Deary, I.J. (2009). Measuring voice outcomes: state of the science review. J Laryngol Otol *123*, 823–829.

Carpenter, S., and Karpati, G. (1982). Necrosis of capillaries in denervation atrophy of human skeletal muscle. Muscle & Nerve *5*, 250–254.

Carroll, T.L., and Rosen, C.A. (2011). Long-term results of calcium hydroxylapatite for vocal fold augmentation. The Laryngoscope *121*, 313–319.

Charles, C., Whelan, T., Gafni, A., Reyno, L., and Redko, C. (1998). Doing Nothing is No Choice: Lay Constructions of Treatment Decision-making Among Women with Early-stage Breast Cancer. Sociology of Health & Illness *20*, 71–95.

Chen, D., Chen, S., Wang, W., Zhang, C., and Zheng, H. (2011). Spontaneous regeneration of recurrent laryngeal nerve following long-term vocal fold paralysis in humans: histologic evidence. Laryngoscope *121*, 1035–1039.

Chen, H.-C., Jen, Y.-M., Wang, C.-H., Lee, J.-C., and Lin, Y.-S. (2007). Etiology of vocal cord paralysis. ORL J. Otorhinolaryngol. Relat. Spec. *69*, 167–171.

Cheung, E.J., and McGinn, J.D. (2007). The surgical treatment of bilateral vocal fold impairment. Operative Techniques in Otolaryngology-Head and Neck Surgery *18*, 144–155.

Chinn, S. (1991). Statistics in respiratory medicine. 2. Repeatability and method comparison. Thorax *46*, 454–456.

Chou, F.-F., Su, C.-Y., Jeng, S.-F., Hsu, K.-L., and Lu, K.-Y. (2003). Neurorrhaphy of the recurrent laryngeal nerve. Journal of the American College of Surgeons *197*, 52–57.

Chowdhury, K., Saha, S., Saha, V.P., Pal, S., and Chatterjee, I. (2013). Pre and Post Operative Voice Analysis After Medialization Thyroplasty in Cases of Unilateral Vocal Fold Paralysis. Indian J Otolaryngol Head Neck Surg 354–357.

Chunying, F., Haifeng, L., Lin, M., and Xiaopeng, Z. (2013). Nonlinear Dynamic Analysis of Pathological Voices. In Intelligent Computing Theories and Technology, D.-S. Huang, K.-H. Jo, Y.-Q. Zhou, and K. Han, eds. (Springer Berlin Heidelberg), pp. 401–409.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychological Bulletin *70*, 213–220.

Cohen, S.M., Pitman, M.J., Noordzij, J.P., and Courey, M. (2012). Evaluation of Dysphonic Patients by General Otolaryngologists. Journal of Voice *26*, 772–778.

Connor, S.E.J., Chaudhary, N., Fareedi, S., and Woo, E.K. (2006). Imaging of muscular denervation secondary to motor cranial nerve dysfunction. Clinical Radiology *61*, 659–669.

Cook, R.C., Alscher, K.T., and Hsiang, Y.N. (2003). A debate on the value and necessity of clinical trials in surgery. The American Journal of Surgery *185*, 305–310.

Crumley, R.L. (1982). Experiments in laryngeal reinnervation. The Laryngoscope *92*, 1–27.

Crumley, R.L. (1983). Phrenic nerve graft for bilateral vocal cord paralysis. Laryngoscope *93*, 425–428.

Crumley, R.L. (1991). Update: ansa cervicalis to recurrent laryngeal nerve anastomosis for unilateral laryngeal paralysis. Laryngoscope *101*, 384–387; discussion 388.

Crumley, R.L. (1994). Unilateral recurrent laryngeal nerve paralysis. Journal of Voice *8*, 79–83.

Crumley, R.L. (2000). Laryngeal synkinesis revisited. Ann. Otol. Rhinol. Laryngol. *109*, 365–371.

Crumley, R.L., and Izdebski, K. (1986). Voice quality following laryngeal reinnervation by ansa hypoglossi transfer. Laryngoscope *96*, 611–616.

Crumley, R.L., Izdebski, K., and McMicken, B. (1988). Nerve transfer versus Teflon® injection for vocal cord paralysis: A comparison. Laryngoscope *98*, 1200–1204.

Cutler, J.L., and Cleveland, T. (2002). The clinical usefulness of laryngeal videostroboscopy and the role of high-speed cinematography in laryngeal evaluation. Current Opinion in Otolaryngology & Head & Neck Surgery December 2002 *10*, 462–466.

Czaja Pawel, Soja Jerzy, Cmiel Adam, Szczeklik Andrzej, and Sladek Krzysztof (2007). Assessment of Airway Caliber in Quantitative Videobronchoscopy. Respiration *74*, 432–438.

Dailey, S.H., Kobler, J.B., Hillman, R.E., Tangrom, K., Thananart, E., Mauri, M., and Zeitels, S.M. (2005). Endoscopic Measurement of Vocal Fold Movement during Adduction and Abduction. The Laryngoscope *115*, 178–183.

Dalgic, A., Kandogan, T., Koc, M., Kulan, C.A., Yagci, A., Engin, O., Aksoy, G., and Ozuer, M.Z. (2013). Short-term laryngeal electromyography and histopathological findings after primary reconstruction of the inferior laryngeal nerve in rabbits: prospective study. The Journal of Laryngology & Otology *127*, 48–53.

Daugherty, C., Ratain, M.J., Grochowski, E., Stocking, C., Kodish, E., Mick, R., and Siegler, M. (1995). Perceptions of cancer patients and their physicians involved in phase I trials. JCO *13*, 1062–1072.

Davis, C.P., Mckinnon, G.C., Debatin, J.F., and Schulthess, G.K. (1996). Ultra-high-speed MR imaging. European Radiology *6*, 297–311.

Deary, I.J., Webb, A., MacKenzie, K., Wilson, J.A., and Carding, P.N. (2004). Short, self-report voice symptom scales: Psychometric characteristics of the voice handicap index-10 and the vocal performance questionnaire. Otolaryngology - Head and Neck Surgery *131*, 232–235.

Dehqan, A., Ansari, H., and Bakhtiar, M. (2010). Objective Voice Analysis of Iranian Speakers with Normal Voices. Journal of Voice *24*, 161–167.

Dejonckere, P.H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Van De Heyning, P., Remacle, M., and Woisard, V. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). Eur Arch Otorhinolaryngol *258*, 77–82.

Dennis, D., and Kashima, H. (1989). Carbon dioxide laser posterior cordectomy for treatment of bilateral vocal cord paralysis. Ann Otol Rhinol Laryngol *98*, 930–934.

Donovan, J. (2002). Quality improvement report: Improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study * Commentary: presenting unbiased information to patients can be difficult. BMJ *325*, 766–770.

Donovan, J.L., Athene Lane, J., Peters, T.J., Brindle, L., Salter, E., Gillatt, D., Powell, P., Bollina, P., Neal, D.E., and Hamdy, F.C. (2009). Development of a complex intervention improved randomization and informed consent in a randomized controlled trial. Journal of Clinical Epidemiology *62*, 29–36.

Donovan, J.L., Paramasivan, S., Salis, I. de, and Toerien, M. (2014). Clear obstacles and hidden challenges: understanding recruiter perspectives in six pragmatic randomised controlled. Trials *15*, 5.

Edwards, S.J.L., Braunholtz, D.A., Lilford, R.J., and Stevens, A.J. (1999). Ethical issues in the design and conduct of cluster randomised controlled trials. BMJ *318*, 1407–1409.

Egginton, S., Hudlická, O., Brown, M.D., Walter, H., Weiss, J.B., and Bate, A. (1998). Capillary growth in relation to blood flow and performance in overloaded rat skeletal muscle. Journal of Applied Physiology *85*, 2025–2032.

El-Kashlan, H.K., Carroll, W.R., Hogikyan, N.D., Chepeha, D.B., Kileny, P.R., and Esclamado, R.M. (2001). Selective cricothyroid muscle reinnervation by muscle-nerve-muscle neurotization. Arch. Otolaryngol. Head Neck Surg *127*, 1211–1215.

Eller, R.L., Miller, M., Weinstein, J., and Sataloff, R.T. (2009). The innervation of the posterior cricoarytenoid muscle: exploring clinical possibilities. J Voice *23*, 229–234.

Ellis, P.M., Butow, P.N., and Tattersall, M.H.N. (2002). Informing breast cancer patients about clinical trials: a randomized clinical trial of an educational booklet. Ann Oncol *13*, 1414–1423.

Ergina, P.L., Cook, J.A., Blazeby, J.M., Boutron, I., Clavien, P.-A., Reeves, B.C., and Seiler, C.M. (2009). Challenges in evaluating surgical innovation. The Lancet *374*, 1097–1104.

Fackler, C.D., Perret, G.E., and Bedell, G.N. (1967). Effect of unilateral phrenic nerve section on lung function. Journal of Applied Physiology *23*, 923–926.

Fang, T.-J., Li, H.-Y., Gliklich, R.E., Chen, Y.-H., Wang, P.-C., and Chuang, H.-F. (2008). Quality of life measures and predictors for adults with unilateral vocal cord paralysis. Laryngoscope *118*, 1837–1841.

Fang, T.-J., Tam, Y.-Y., Courey, M.S., Li, H.-Y., and Chiang, H.-C. (2011). Unilateral high vagal paralysis: relationship of the severity of swallowing disturbance and types of injuries. Laryngoscope *121*, 245–249.

Faure, M.-A., and Muller, A. (1992). Stroboscopy. Journal of Voice *6*, 139–148.

Faust, R.A., Remley, K.B., and Rimell, F.L. (2001). Real-Time, Cine Magnetic Resonance Imaging for Evaluation of the Pediatric Airway. The Laryngoscope *111*, 2187–2190.

Franco Jr, R.A. (2012). Adduction arytenopexy, hypopharyngoplasty, medialization laryngoplasty, and cricothyroid subluxation for the treatment of paralytic dysphonia and dysphagia. Operative Techniques in Otolaryngology *23*, 164-172.

Featherstone, K., and Donovan, J.L. (2002). "Why don't they just tell me straight, why allocate it?" The struggle to make sense of participating in a randomised controlled trial. Social Science & Medicine *55*, 709–719.

Felippe, A.C.N. de, Grillo, M.H.M.M., and Grechi, T.H. (2006). Standardization of acoustic measures for normal voice patterns. Revista Brasileira de Otorrinolaringologia *72*, 659–664.

Fex, S. (1970). Functioning remobilization of vocal cords in cats with permanent recurrent laryngeal nerve paresis. Acta Otolaryngol *69*, 294–301.

Flint, P.W., Downs, D.H., and Coltrera, M.D. (1991). Laryngeal synkinesis following reinnervation in the rat. Neuroanatomic and physiologic study using retrograde fluorescent tracers and electromyography. Ann. Otol. Rhinol. Laryngol *100*, 797–806.

Flint, P.W., Purcell, L.L., and Cummings, C.W. (1997). Pathophysiology and indications for medialization thyroplasty in patients with dysphagia and aspiration. Otolaryngol Head Neck Surg *116*, 349–354.

Forkert, L., Watanabe, H., Sutherland, K., Vincent, S., and Fisher, J.T. (1996). Quantitative videobronchoscopy: a new technique to assess airway caliber. Am. J. Respir. Crit. Care Med. *154*, 1794–1803.

Gacek, R.R. (2001). Morphologic Correlates for Laryngeal Reinnervation. The Laryngoscope *111*, 1871–1877.

Gilmer-Hill, H.S., and Kline, D.G. (2000). Neurogenic tumors of the cervical vagus nerve: report of four cases and review of the literature. Neurosurgery *46*, 1498–1503.

Gliklich, R.E., Glovsky, R.M., and Montgomery, W.W. (1999). Validation of a voice outcome survey for unilateral vocal cord paralysis. Otolaryngol Head Neck Surg *120*, 153–158.

Godino-Llorente, J.I., Osma-Ruiz, V., Sáenz-Lechón, N., Cobeta-Marco, I., González-Herranz, R., and Ramírez-Calvo, C. (2008). Acoustic analysis of voice

using WPCVox: a comparative study with Multi Dimensional Voice Program. Eur Arch Otorhinolaryngol *265*, 465–476.

Gorphe, P., Hartl, D., Primov-Fever, A., Hans, S., Crevier-Buchman, L., and Brasnu, D. (2013). Endoscopic laser medial arytenoidectomy for treatment of bilateral vocal fold paralysis. European Archives of Oto-Rhino-Laryngology : Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery *270*, 1701–1705.

Goyault, G., Bierry, G., Holl, N., Lhermitte, B., Dietemann, J.L., Beregi, J.P., and Kremer, S. (2012). Diffusion-weighted MRI, dynamic susceptibility contrast MRI and ultrasound perfusion quantification of denervated muscle in rabbits. Skeletal Radiol. *41*, 33–40.

Hamilton, D.W., Salis, I. de, Donovan, J.L., and Birchall, M. (2013). The recruitment of patients to trials in head and neck cancer: a qualitative study of the EaStER trial of treatments for early laryngeal cancer. Eur Arch Otorhinolaryngol *270*, 2333–2337.

Hamy, V., Dikaios, N., Punwani, S., Melbourne, A., Latifoltojar, A., Makanyanga, J., Chouhan, M., Helbren, E., Menys, A., Taylor, S., et al. (2014). Respiratory motion correction in dynamic MRI using robust data decomposition registration – Application to DCE-MRI. Medical Image Analysis *18*, 301–313.

Harries, M.L., and Morrison, M. (1996). The Role of Stroboscopy in the Management of a Patient with a Unilateral Vocal Fold Paralysis. The Journal of Laryngology & Otology *110*, 141–143.

Hartl, D.M., Travagli, J.-P., Leboulleux, S., Baudin, E., Brasnu, D.F., and Schlumberger, M. (2005). Current Concepts in the Management of Unilateral Recurrent Laryngeal Nerve Paralysis after Thyroid Surgery. JCEM *90*, 3084–3088.

Hassan, M.M., Yumoto, E., Baraka, M.A., Sanuki, T., and Kodama, N. (2011). Arytenoid rotation and nerve-muscle pedicle transfer in paralytic dysphonia. Laryngoscope *121*, 1018–1022.

Havas, T., Lowinger, D., and Priestley, J. (1999). Unilateral Vocal Fold Paralysis: Causes, Options and Outcomes. Aust. N.Z. J. Surg. *69*, 509–513.

Hertegård, S., Hallén, L., Laurent, C., Lindström, E., Olofsson, K., Testad, P., and Dahlqvist, Å. (2004). Cross-linked hyaluronan versus collagen for injection treatment of glottal insufficiency: 2-year follow-up. Acta Otolaryngol *124*, 1208–1214.

Hillel, A.D. (2001). The Study of Laryngeal Muscle Activity in Normal Human Subjects and in Patients With Laryngeal Dystonia Using Multiple Fine-Wire Electromyography. The Laryngoscope *111*, 1–47.

Hirano, M. (1974). Morphological Structure of the Vocal Cord as a Vibrator and its Variations. Folia Phoniatrica et Logopaedica *26*, 89–94.

Hirano, M. (1981). Clinical examination of voice (Springer-Verlag).

Hirano, M., and Bless, D.M. (1993). Videostroboscopic Examination of the Larynx (London: Whurr).

Hogikyan, N.D., and Rosen, C.A. (2002). A review of outcome measurements for voice disorders. Otolaryngology - Head and Neck Surgery *126*, 562–572.

Hogikyan, N.D., and Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). Journal of Voice *13*, 557–569.

Hogikyan, N.D., Wodchis, W.P., Terrell, J.E., Bradford, C.R., and Esclamado, R.M. (2000). Voice-Related Quality of Life (V-RQOL) following type I thyroplasty for unilateral vocal fold paralysis. Journal of Voice *14*, 378–386.

Holl, N., Echaniz-Laguna, A., Bierry, G., Mohr, M., Loeffler, J.-P., Moser, T., Dietemann, J.-L., and Kremer, S. (2008). Diffusion-weighted MRI of denervated muscle: a clinical and experimental study. Skeletal Radiol. *37*, 1111–1117.

Hornibrook, J., Lin, E., and Ormond, T. (2011). Using the iPhone for voice recording in laryngology. Otolaryngol Head Neck Surg. *145*, 158.

Hornibrook, J., Lin, E., and Ormond, T. (2012). Evaluating iPhone recordings for acoustic voice assessment. Folia Phoniatr Logop. *64,* 122-130.

Horsthuis, K., Lavini, C., Bipat, S., Stokkers, P.C.F., and Stoker, J. (2009). Perianal Crohn Disease: Evaluation of Dynamic Contrast-enhanced MR Imaging as an Indicator of Disease Activity1. Radiology *251*, 380–387.

Hydman, J., and Mattsson, P. (2008). Collateral reinnervation by the superior laryngeal nerve after recurrent laryngeal nerve injury. Muscle Nerve *38*, 1280–1289.

Inagi, K., Khidr, A.A., Ford, C.N., Bless, D.M., and Heisey, D.M. (1997a). Correlation Between Vocal Functions and Glottal Measurements in Patients With Unilateral Vocal Fold Paralysis. The Laryngoscope *107*, 782–791.

Inagi, K., Khidr, A.A., Ford, C.N., Bless, D.M., and Heisey, D.M. (1997b). Correlation Between Vocal Functions and Glottal Measurements in Patients With Unilateral Vocal Fold Paralysis. The Laryngoscope *107*, 782–791.

Isshiki, N., Okamura, H., and Ishikawa, T. (1975). Thyroplasty type I (lateral compression) for dysphonia due to vocal cord paralysis or atrophy. Acta Oto-Laryngologica *80*, 465–473.

Isshiki, N., Tanabe, M., and Sawada, M. (1978). Arytenoid adduction for unilateral vocal cord paralysis. Archives of Otolaryngology *104*, 555–558.

Jacobson, B.H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M.S., and Newman, C.W. (1997). The Voice Handicap Index (VHI): Development and Validation. Am J Speech Lang Pathol *6*, 66–70.

Jeannon, J.-P., Orabi, A.A., Bruch, G.A., Abdalsalam, H.A., and Simo, R. (2009). Diagnosis of recurrent laryngeal nerve palsy after thyroidectomy: a systematic review. Int. J. Clin. Pract *63*, 624–629.

Jiang, H., Wang, Z., Xian, J., Li, J., Chen, Q., and Ai, L. (2012). Evaluation of rectus extraocular muscles using dynamic contrast-enhanced MR imaging in patients with Graves' ophthalmopathy for assessment of disease activity. Acta Radiol *53*, 87–94.

Kamath, S., Venkatanarasimha, N., Walsh, M.A., and Hughes, P.M. (2007). MRI appearance of muscle denervation. Skeletal Radiol *37*, 397–404.

Kano, S., Horowitz, J.B., and Sasaki, C.T. (1991). Posterior cricoarytenoid muscle denervation. Arch. Otolaryngol. Head Neck Surg. *117*, 1019–1020.

Karnell, M.P., Hall, K.D., and Landahl, K.L. (1995). Comparison of fundamental frequency and perturbation measurements among three analysis systems. Journal of Voice *9*, 383–393.

Kato, K., Tomura, N., Takahashi, S., and Watarai, J. (2002). Motor Denervation of Tumors of the Head and Neck: Changes in MR Appearance. Magnetic Resonance in Medical Sciences *1*, 157–164.

Khoo, M.M.Y., Tyler, P.A., Saifuddin, A., and Padhani, A.R. (2011). Diffusion-weighted imaging (DWI) in musculoskeletal MRI: a critical review. Skeletal Radiol. *40*, 665–681.

Kikuchi, Y., Nakamura, T., Takayama, S., Horiuchi, Y., and Toyama, Y. (2003). MR imaging in the diagnosis of denervated and reinnervated skeletal muscles: experimental study in rats. Radiology *229*, 861–867.

Kingham, P.J., Birchall, M.A., Burt, R., Jones, A., and Terenghi, G. (2005). Reinnervation of laryngeal muscles: A study of changes in myosin heavy chain expression. Muscle & Nerve *32*, 761–766.

Klingholtz, F. (1990). Acoustic recognition of voice disorders: A comparative study of running speech versus sustained vowels. The Journal of the Acoustical Society of America *87*, 2218–2224.

Knopp, M.V.M.D., Giesel, F.L., Marcos, H., von Tengg-Kobligk, H., and Choyke, P. (2001). Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Oncology. Topics in Magnetic Resonance Imaging August 2001 *12*, 301–308.

Knopp, Michael V, Giesel, Frederik L, Marcos, Hani, von Tengg-Kobligk, Hendrik, and Choyke, Peter Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Onco... : Topics in Magnetic Resonance Imaging.

Koltzenburg, M., and Yousry, T. (2007). Magnetic resonance imaging of skeletal muscle. Current Opinion in Neurology *20*, 595–599.

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B.J., Hróbjartsson, A., Roberts, C., Shoukri, M., and Streiner, D.L. (2011). Guidelines for Reporting

Reliability and Agreement Studies (GRRAS) were proposed. International Journal of Nursing Studies *48*, 661–671.

Koufman, J.A., and Walker, F.O. (1998). Laryngeal electromyography in clinical practice: indications, techniques, and interpretation. Phonoscope *1*, 57–70.

Koufman, J.A., Postma, G.N., Whang, C.S., Rees, C.J., Amin, M.R., Belafsky, P.C., Johnson, P.E., Connolly, K.M., and Walker, F.O. (2001). Diagnostic laryngeal electromyography: The wake forest experience 1995-1999. Otolaryngology - Head and Neck Surgery *124*, 603–606.

Krom, G. de (1995). Some Spectral Correlates of Pathological Breathy and Rough Voice Quality for Different Types of Vowel Fragments. Journal of Speech Language and Hearing Research *38*, 794.

Küllmer, K., Sievers, K.W., Reimers, C.D., Rompe, J.D., Müller-Felber, W., Nägele, M., and Harland, U. (1998). Changes of sonographic, magnetic resonance tomographic, electromyographic, and histopathologic findings within a 2-month period of examinations after experimental muscle denervation. Arch Orthop Trauma Surg *117*, 228–234.

Kwak, P.E., Friedman, A.D., Lamarre, E.D., and Lorenz, R.R. (2010). Selective reinnervation of the posterior cricoarytenoid and interarytenoid muscles: an anatomical study. Laryngoscope *120*, 463–467.

Kwon, T.-K., and Buckmire, R. (2004). Injection laryngoplasty for management of unilateral vocal fold paralysis. Curr Opin Otolaryngol Head Neck Surg *12*, 538–542.

Lakhani, R., Fishman, J.M., Bleach, N., Costello, D., and Birchall, M. (1996). Alternative injectable materials for vocal fold medialisation in unilateral vocal fold paralysis. In Cochrane Database of Systematic Reviews, (John Wiley & Sons, Ltd),.

Lamarre, E.D., Lorenz, R.R., Milstein, C., and Scharpf, J. (2010). Laryngeal reinnervation after vagal paraganglioma resection: a case report. Am J Otolaryngol.

Lamarre, E.D., Lorenz, R.R., Milstein, C., and Scharpf, J. (2011). Laryngeal reinnervation after vagal paraganglioma resection: a case report. Am J Otolaryngol *32*, 171–173.

Landis, J.R., and Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics *33*, 159.

Langer, D.L., van der Kwast, T.H., Evans, A.J., Trachtenberg, J., Wilson, B.C., and Haider, M.A. (2009). Prostate cancer detection with multi-parametric MRI: Logistic regression analysis of quantitative T2, diffusion-weighted imaging, and dynamic contrast-enhanced MRI. J. Magn. Reson. Imaging *30*, 327–334.

Lau, D.P., Lee, G.A., Wong, S.M., Lim, V.P., Chan, Y.H., Tan, N.G., Rammage, L.A., and Morrison, M.D. (2010). Injection Laryngoplasty With Hyaluronic Acid for Unilateral Vocal Cord Paralysis. Randomized Controlled Trial Comparing Two Different Particle Sizes. Journal of Voice *24*, 113–118.

Lavini, C., Buiter, and Maas (2013). Use of dynamic contrast enhanced time intensity curve shape analysis in MRI: theory and practice. Reports in Medical Imaging 71.

Leão, H.Z., de Lima Pontes, P.A., Rosito Coiro, J.R., and Lamberts, M. (2010). Ultra Structure of the Denervated Vocal Muscle Mechanically in Hogs. Intl. Arch. Otorhinolaryngol. *14*, 66–81.

Leder, S.B., and Sasaki, C.T. (1994). Long-Term Changes in Vocal Quality Following Isshiki Thyroplasty Type I. The Laryngoscope *104*, 275–277.

Lee, J., Koh, D., and Ong, C.N. (1989). Statistical evaluation of agreement between two methods for measuring a quantitative variable. Computers in Biology and Medicine *19*, 61–70.

Lee, W.T., Milstein, C., Hicks, D., Akst, L.M., and Esclamado, R.M. (2007a). Results of ansa to recurrent laryngeal nerve reinnervation. Otolaryngology - Head and Neck Surgery *136*, 450–454.

Lee, W.T., Milstein, C., Hicks, D., Akst, L.M., and Esclamado, R.M. (2007b). Results of ansa to recurrent laryngeal nerve reinnervation. Otolaryngol Head Neck Surg *136*, 450–454.

Lewin, S., Glenton, C., and Oxman, A.D. (2009). Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. BMJ *339*, b3496.

Lewis, W.S., Crumley, R.L., Pitcock, J.K., and Blanks, R.H.I. (1991). Does intralaryngeal motor nerve sprouting occur following unilateral recurrent laryngeal nerve paralysis? The Laryngoscope *101*, 1259–1263.

Li, M., Liu, F., Shi, S., Chen, S., Chen, D., and Zheng, H. (2013). Bridging Gaps Between the Recurrent Laryngeal Nerve and Ansa Cervicalis Using Autologous Nerve Grafts. Journal of Voice *27*, 381–387.

Lilford, R., Braunholtz, D., Harris, J., and Gill, T. (2004). Trials in surgery. Br J Surg *91*, 6–16.

Litton, E., and Delaney, A. (2013). Prophylactic intraaortic balloon counterpulsation in high-risk cardiac surgery: a survey of opinion and current practice. HSR Proc Intensive Care Cardiovasc Anesth *5*, 33–39.

Lloyd, P.G., Prior, B.M., Yang, H.T., and Terjung, R.L. (2003). Angiogenic growth factor expression in rat skeletal muscle in response to exercise training. Am J Physiol Heart Circ Physiol *284*, H1668–H1678.

Lorenz, R.R., Esclamado, R.M., Teker, A.M., Strome, M., Scharpf, J., Hicks, D., Milstein, C., and Lee, W.T. (2008). Ansa cervicalis-to-recurrent laryngeal nerve anastomosis for unilateral vocal fold paralysis: experience of a single institution. Ann. Otol. Rhinol. Laryngol *117*, 40–45.

Lovato, L.C., Hill, K., Hertert, S., Hunninghake, D.B., and Probstfield, J.L. (1997). Recruitment for controlled clinical trials: Literature summary and annotated bibliography. Controlled Clinical Trials *18*, 328–352.

Lozano, F.S., Gómez, J.L., Mondillo, M.C., González-Porras, J.R., González-Sarmiento, R., and Muñoz, A. (2008). Surgery of vagal paragangliomas: Six patients and review of literature. Surgical Oncology *17*, 281–287.

Lu, D.-X., Huang, S.-K., and Carlson, B.M. (1997). Electron microscopic study of long-term denervated rat skeletal muscle. The Anatomical Record *248*, 355–365.

Ma, E.P.-M., and Yiu, E.M.-L. (2005). Suitability of Acoustic Perturbation Measures in Analysing Periodic and Nearly Periodic Voice Signals. Folia Phoniatrica et Logopaedica *57*, 38–47.

Mallur, P.S., and Rosen, C.A. (2010). Vocal fold injection: review of indications, techniques, and materials for augmentation. Clin Exp Otorhinolaryngol *3*, 177–182.

Maranillo, E., León, X., Ibañez, M., Orús, C., Quer, M., and Sañudo, J.R. (2003). Variability of the nerve supply patterns of the human posterior cricoarytenoid muscle. Laryngoscope *113*, 602–606.

Marcum, K.K., Wright Jr., S.C., Kemp, E.S., and Kitse, D.J. (2010). A novel modification of the ansa to recurrent laryngeal nerve reinnervation procedure for young children. International Journal of Pediatric Otorhinolaryngology *74*, 1335–1337.

Marie, J.P., Dehesdin, D., Ducastelle, T., and Senant, J. (1989). Selective reinnervation of the abductor and adductor muscles of the canine larynx after recurrent nerve paralysis. Ann. Otol. Rhinol. Laryngol *98*, 530–536.

Marie, J.P., Tardif, C., Lerosey, Y., Gibon, J.F., Hellot, M.F., Tadié, M., Andrieu-Guitrancourt, J., Dehesdin, D., and Pasquis, P. (1997). Selective resection of the phrenic nerve roots in rabbits: Part II: Respiratory effects. Respiration Physiology *109*, 139–148.

Marina, M.B., Marie, J.-P., and Birchall, M.A. (2011). Laryngeal reinnervation for bilateral vocal fold paralysis. Curr Opin Otolaryngol Head Neck Surg *19*, 434–438.

Maronian, N., Waugh, P., Robinson, L., and Hillel, A. (2003). Electromyographic findings in recurrent laryngeal nerve reinnervation. Ann. Otol. Rhinol. Laryngol *112*, 314–323.

Maronian, N.C., Robinson, L., Waugh, P., and Hillel, A.D. (2004). A new electromyographic definition of laryngeal synkinesis. Ann. Otol. Rhinol. Laryngol *113*, 877–886.

Martin-Oviedo, C., Maranillo, E., Lowy-Benoliel, A., Pascual-Font, A., Martinez-Guirado, T., Rodriguez-Niedenführ, M., Sañudo, J., Scola, B., and Vazquez, T. (2011). Functional role of human laryngeal nerve connections. The Laryngoscope *121*, 2338–2343.

Maryn, Y., Roy, N., Bodt, M.D., Cauwenberge, P.V., and Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. The Journal of the Acoustical Society of America *126*, 2619–2634.

May, M., and Beery, Q. (1986). Muscle-nerve pedicle laryngeal reinnervation. Laryngoscope *96*, 1196–1200.

McCulloch, P. (2002). Randomised trials in surgery: problems and possible solutions. BMJ *324*, 1448–1451.

McCulloch, T.M., and Hoffman, H.T. (1998). Medialization laryngoplasty with expanded polytetrafluoroethylene. Surgical technique and preliminary results. Ann. Otol. Rhinol. Laryngol *107*, 427–432.

McCulloch, T.M., Andrews, B.T., Hoffman, H.T., Graham, S.M., Karnell, M.P., and Minnick, C. (2002). Long-Term Follow-up of Fat Injection Laryngoplasty for Unilateral Vocal Cord Paralysis. The Laryngoscope *112*, 1235–1238.

McLean-Muse, A., Montgomery, W.W., Hillman, R.E., Varvares, M., Bunting, G., Doyle, P., and Eng, J. (2000). Montgomery Thyroplasty Implant for vocal fold immobility: phonatory outcomes. Ann. Otol. Rhinol. Laryngol *109*, 393–400.

McLeod, R.S. (1999). Issues in surgical randomized controlled trials. World J Surg *23*, 1210–1214.

Md, A.R., Md, L.S., and Osman Ratib MD, P. (2004). OsiriX: An Open-Source Software for Navigating in Multidimensional DICOM Images. J Digit Imaging *17*, 205–216.

Miller, R.B., Boon, M.S., Atkins, J.P., and Lowry, L.D. (2000). Vagal paraganglioma: The Jefferson experience. Otolaryngology - Head and Neck Surgery *122*, 482–487.

Mills, N., Donovan, J.L., Smith, M., Jacoby, A., Neal, D.E., and Hamdy, F.C. (2003). Perceptions of equipoise are crucial to trial participation: a qualitative study of men in the ProtecT study. Controlled Clinical Trials *24*, 272–282.

Misono, S., and Merati, A.L. (2012). Evidence-based practice: evaluation and management of unilateral vocal fold paralysis. Otolaryngol. Clin. North Am. *45*, 1083–1108.

Miyamaru, S., Kumai, Y., Ito, T., and Yumoto, E. (2008). Effects of Long-Term Denervation on the Rat Thyroarytenoid Muscle. The Laryngoscope *118*, 1318–1323.

Miyauchi, A., Inoue, H., Tomoda, C., Fukushima, M., Kihara, M., Higashiyama, T., Takamura, Y., Ito, Y., Kobayashi, K., and Miya, A. (2009). Improvement in phonation after reconstruction of the recurrent laryngeal nerve in patients with thyroid cancer invading the nerve. Surgery *146*, 1056–1062.

Montgomery, W.W., and Montgomery, S.K. (1997). Montgomery thyroplasty implant system. Ann Otol Rhinol Laryngol Suppl *170*, 1–16.

Moon, K.R., Chung, S.M., Park, H.S., and Kim, H.S. (2012). Materials of Acoustic Analysis: Sustained Vowel Versus Sentence. Journal of Voice *26*, 563–565.

Moonis, G., Dyce, O., Loevner, L.A., and Mirza, N. (2005). Magnetic resonance imaging of micronized dermal graft in the larynx. Ann. Otol. Rhinol. Laryngol. *114*, 593–598.

Morgan, J.E., Zraick, R.I., Griffin, A.W., Bowen, T.L., and Johnson, F.L. (2007). Injection Versus Medialization Laryngoplasty for the Treatment of Unilateral Vocal Fold Paralysis. The Laryngoscope *117*, 2068–2074.

Mossallam, I., Nasser Kotby, M., Abd-El-Rahman, S., and El-Samma, M. (1987). Attachment of some internal laryngeal muscles at the base of the arytenoid cartilage. Acta Oto-Laryngologica *103*, 649–656.

Mostafa, B.E., Gadallah, N.A., Nassar, N.M., Al Ibiary, H.M., Fahmy, H.A., and Fouda, N.M. (2004). The Role of Laryngeal Electromyography in Vocal Fold Immobility. ORL *66*, 5–10.

Motoyoshi, K., Hyodo, M., Yamagata, T., and Gyo, K. (2004). Restoring Vocal Fold Movement after Transection and Immediate Suturing of the Recurrent Laryngeal Nerve with Local Application of Basic Fibroblast Growth Factor: An Experimental Study in the Rat. The Laryngoscope *114*, 1247–1252.

Munin, M.C., Rosen, C.A., and Zullo, T. (2003). Utility of laryngeal electromyography in predicting recovery after vocal fold paralysis. Archives of Physical Medicine and Rehabilitation *84*, 1150–1153.

Murakami, R., Baba, Y., Nishimura, R., Baba, T., Okuda, T., Utsunomiya, D., Ishikawa, T., and Takahashi, M. (1998). MR of denervated tongue: temporal changes after radical neck dissection. AJNR Am J Neuroradiol *19*, 515–518.

Murtagh, M.J., Thomson, R.G., May, C.R., Rapley, T., Heaven, B.R., Graham, R.H., Kaner, E.F., Stobbart, L., and Eccles, M.P. (2007). Qualitative methods in a randomised controlled trial: the role of an integrated qualitative process evaluation in providing evidence to discontinue the intervention in one arm of a trial of a decision support tool. Qual Saf Health Care *16*, 224–229.

Nasri, S., Beizai, P., Ye, Y.M., Sercaz, J.A., Kim, Y.M., and Berke, G.S. (1997). Cross-innervation of the thyroarytenoid muscle by a branch from the external division of the superior laryngeal nerve. Ann. Otol. Rhinol. Laryngol. *106*, 594–598.

Nawka, T., and Konerding, U. (2012). The Interrater Reliability of Stroboscopy Evaluations. Journal of Voice *26*, 812.e1–e812.e10.

Netterville, J.L., Jackson, C.G., and Civantos, F. (1993). Thyroplasty in the functional rehabilitation of neurotologic skull base surgery patients. Am J Otol *14*, 460–464.

Netterville JL, Jackson C, Miller FR, Wanamaker JR, and Glasscock ME (1998). Vagal paraganglioma: A review of 46 patients treated during a 20-year period. Arch Otolaryngol Head Neck Surg *124*, 1133–1140.

Olson, D.E., Goding, G.S., and Michael, D.D. (1998). Acoustic and perceptual evaluation of laryngeal reinnervation by ansa cervicalis transfer. Laryngoscope *108*, 1767–1772.

Olthoff, A., Zeiss, D., Laskawi, R., Kruse, E., and Steiner, W. (2005). Laser Microsurgical Bilateral Posterior Cordectomy for the Treatment of Bilateral Vocal Fold Paralysis. The Annals of Otology, Rhinology & Laryngology *114*, 599–604.

Omori, K., Kacker, A., Slavit, D.H., and Blaugrund, S.M. (1996). Quantitative videostroboscopic measurement of glottal gap and vocal function: an analysis of thyroplasty type I. Ann. Otol. Rhinol. Laryngol. *105*, 280–285.

Padhani, A.R., Liu, G., Mu-Koh, D., Chenevert, T.L., Thoeny, H.C., Takahara, T., Dzik-Jurasz, A., Ross, B.D., Van Cauteren, M., Collins, D., et al. (2009). Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. Neoplasia *11*, 102–125.

Palmer, A.J.R., Thomas, G.E.R., Pollard, T.C.B., Rombach, I., Taylor, A., Arden, N., Beard, D.J., Andrade, A.J., Carr, A.J., and Glyn-Jones, S. (2013). The feasibility of performing a randomised controlled trial for femoroacetabular impingement surgery. Bone Joint Res *2*, 33–40.

Paniello, R.C. (2000). Laryngeal reinnervation with the hypoglossal nerve: II. Clinical evaluation and early patient experience. Laryngoscope *110*, 739–748.

Paniello, R.C., Edgar, J.D., Kallogjeri, D., and Piccirillo, J.F. (2011). Medialization versus reinnervation for unilateral vocal fold paralysis: A multicenter randomized clinical trial. Laryngoscope *121*, 2172–2179.

Parsa, V., and Jamieson, D.G. (2001). Acoustic Discrimination of Pathological Voice: Sustained Vowels Versus Continuous Speech. Journal of Speech Language and Hearing Research *44*, 327.

Paskhover, B., Wadie, M., and Sasaki, C.T. (2014). Thyroarytenoid cross-innervation by the external branch of the superior laryngeal nerve in the porcine model. The Laryngoscope n/a – n/a.

Pitman, M.J., Weissbrod, P., Roark, R., Sharma, S., and Schaefer, S.D. (2011). Electromyographic and histologic evolution of the recurrent laryngeal nerve from transection and anastomosis to mature reinnervation. Laryngoscope *121*, 325–331.

Poburka, B.J. (1999). A new stroboscopy rating form. Journal of Voice *13*, 403–413.

Poburka, B.J., and Bless, D.M. (1998). A multi-media, computer-based method for stroboscopy rating training. J Voice *12*, 513–526.

Polak, J.F., Jolesz, F.A., and Adams, D.F. (1988). Magnetic resonance imaging of skeletal muscle. Prolongation of T1 and T2 subsequent to denervation. Invest Radiol *23*, 365–369.

Portney, M., and Watkins, M.A. (1993). Foundations of clinical research; applications to practice (Prentice Hall Ltd, Sydney).

Punwani, S. (2011). Contrast enhanced MR imaging of female pelvic cancers: Established methods and emerging applications. European Journal of Radiology *78*, 2–11.

Punwani, S., Taylor, S.A., Saad, Z.Z., Bainbridge, A., Groves, A., Daw, S., Shankar, A., Halligan, S., and Humphries, P.D. (2013). Diffusion-weighted MRI of lymphoma: prognostic utility and implications for PET/MRI? Eur J Nucl Med Mol Imaging *40*, 373–385.

Rabinov, C.R., Kreiman, J., Gerratt, B.R., and Bielamowicz, S. (1995). Comparing Reliability of Perceptual Ratings of Roughness and Acoustic Measures of Jitter. J Speech Hear Res *38*, 26–32.

Ramadan, D.H.H., Wax, M.K., and Avery, S. (1998). Outcome and changing cause of unilateral vocal cord paralysis. Otolaryngology - Head and Neck Surgery *118*, 199–202.

Remacle, M., and Eckel, H.E. (2010). Surgery of Larynx and Trachea (Heidelberg: Springer).

Rice, D.H. (1982). Laryngeal Reinnervation. The Laryngoscope *92*, 1049???1059.

Rickert, S.M., Childs, L.F., Carey, B.T., Murry, T., and Sulica, L. (2012). Laryngeal electromyography for prognosis of vocal fold palsy: a meta-analysis. Laryngoscope *122*, 158–161.

Rihkanen, H., Reijonen, P., Lehikoinen-Söderlund, S., and Lauri, E.-R. (2004). Videostroboscopic assessment of unilateral vocal fold paralysis after augmentation with autologous fascia. European Archives of Oto-Rhino-Laryngology *261*, 177–183.

Romo, L.V., and Curtin, H.D. (1999). Atrophy of the posterior cricoarytenoid muscle as an indicator of recurrent laryngeal nerve palsy. AJNR Am J Neuroradiol *20*, 467–471.

Rosen, C.A. (2005). Stroboscopy as a Research Instrument: Development of a Perceptual Evaluation Tool. The Laryngoscope *115*, 423–428.

Rosen, C.A., Murry, T., Zinn, A., Zullo, T., and Sonbolian, M. (2000). Voice handicap index change following treatment of voice disorders. Journal of Voice *14*, 619–623.

Rosen, C.A., Lee, A.S., Osborne, J., Zullo, T., and Murry, T. (2004). Development and validation of the voice handicap index-10. Laryngoscope *114*, 1549–1556.

Rosen, C.A., Gartner-Schmidt, J., Casiano, R., Anderson, T.D., Johnson, F., Reussner, L., Remacle, M., Sataloff, R.T., Abitbol, J., Shaw, G., et al. (2007). Vocal fold augmentation with calcium hydroxylapatite (CaHA). Otolaryngology - Head and Neck Surgery *136*, 198.e1–e198.e12.

Rosen, C.A., Gartner-Schmidt, J., Casiano, R., Anderson, T.D., Johnson, F., Remacle, M., Sataloff, R.T., Abitbol, J., Shaw, G., Archer, S., et al. (2009). Vocal fold augmentation with calcium hydroxylapatite: Twelve-month report. The Laryngoscope *119*, 1033–1041.

Rosenthal, L.H.S., Benninger, M.S., and Deeb, R.H. (2007). Vocal fold immobility: a longitudinal analysis of etiology over 20 years. Laryngoscope *117*, 1864–1870.

Rosow, D.E., and Sulica, L. (2010). Laryngoscopy of vocal fold paralysis: Evaluation of consistency of clinical findings. The Laryngoscope *120*, 1376–1382.

Ross, S., Grant, A., Counsell, C., Gillespie, W., Russell, I., and Prescott, R. (1999). Barriers to Participation in Randomised Controlled Trials: A Systematic Review. Journal of Clinical Epidemiology *52*, 1143–1156.

Rosset, A., Spadola, L., and Ratib, O. (2004). OsiriX: An Open-Source Software for Navigating in Multidimensional DICOM Images. J Digit Imaging *17*, 205–216.

Russo, C.P., Smoker, W.R., and Weissman, J.L. (1997). MR appearance of trigeminal and hypoglossal motor denervation. AJNR Am J Neuroradiol *18*, 1375–1383.

Ryu, I.S., Nam, S.Y., Han, M.W., Choi, S.-H., Kim, S.Y., and Roh, J.-L. (2012). Long-term Voice Outcomes After Thyroplasty for Unilateral Vocal Fold Paralysis. Archives of Otolaryngology--Head & Neck Surgery.

Sakai, F., Gamsu, G., Dillon, W.P., Lynch, D.A., and Gilbert, T.J. (1990). MR imaging of the larynx at 1.5 T. J Comput Assist Tomogr *14*, 60–71.

Schlamann, M., Lehnerdt, G., Maderwald, S., and Ladd, S. (2009). Dynamic MRI of the vocal cords using phased-array coils: A feasibility study. Indian J Radiol Imaging *19*, 127–131.

Schneider-Stickler, B., Gaechter, J., and Bigenzahn, W. (2013). Long-term results after external vocal fold medialization thyroplasty with titanium vocal fold medialization implant (TVFMI). Eur Arch Otorhinolaryngol.

Schwartz, C.E., and Fox, B.H. (1995). Who says yes? Identifying selection biases in a psychosocial intervention study of multiple sclerosis. Social Science & Medicine *40*, 359–370.

Seddon, H.J. (1942). A Classification Of Nerve Injuries. The British Medical Journal *2*, 237–239.

Segas, J., Stavroulakis, P., Manolopoulos, L., Yiotakis, J., and Adamopoulos, G. (2001). Management of bilateral vocal fold paralysis: Experience at the University of Athens. Otolaryngology - Head and Neck Surgery *124*, 68–71.

Sercarz, J.A. (1992). Videostroboscopy of human vocal fold paralysis. Ann. Otol. Rhinol. Laryngol *101*, 567.

Shindo, M.L., Herzon, G.D., Hanson, D.G., Cain, D.J., and Sahgal, V. (1992). Effects of Denervation on Laryngeal Muscles: A Canine Model. The Laryngoscope *102*, 663–669.

Shrout, P.E., and Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin *86*, 420–428.

Smith, E., Gray, S., Verdolini, K., and Lemke, J. (1995). Effects of Voice Disorders on Quality of Life. Otolaryngology -- Head and Neck Surgery *113*, P121.

Smith, L.J., Rosen, C.A., Niyonkuru, C., and Munin, M.C. (2012). Quantitative electromyography improves prediction in vocal fold paralysis. The Laryngoscope *122*, 854–859.

Smith, M.E., Roy, N., and Stoddard, K. (2008). Ansa-RLN reinnervation for unilateral vocal fold paralysis in adolescents and young adults. International Journal of Pediatric Otorhinolaryngology *72*, 1311–1316.

Smits, I., Ceuppens, P., and De Bodt, M.S. (2005). A comparative study of acoustic voice measurements by means of Dr. Speech and Computerized Speech Lab. J Voice *19*, 187–196.

Solomon, M.J., and McLeod, R.S. (1995). Should we be performing more randomized controlled trials evaluating surgical operations? Surgery *118*, 459–467.

Spector, B.C., Netterville, J.L., Billante, C., Clary, J., Reinisch, L., and Smith, T.L. (2001). Quality-of-life assessment in patients with unilateral vocal cord paralysis. Otolaryngology - Head and Neck Surgery *125*, 176–182.

Statham, M.M., Rosen, C.A., Nandedkar, S.D., and Munin, M.C. (2010). Quantitative laryngeal electromyography: Turns and amplitude analysis. The Laryngoscope *120*, 2036–2041.

Su, W.-F., Hsu, Y.-D., Chen, H.-C., and Sheng, H. (2007). Laryngeal reinnervation by ansa cervicalis nerve implantation for unilateral vocal cord paralysis in humans. J. Am. Coll. Surg *204*, 64–72.

Sulica, L. (2008). The Natural History of Idiopathic Unilateral Vocal Fold Paralysis: Evidence and Problems. The Laryngoscope *118*, 1303–1307.

Sunderland, S. (1951). A Classification of Peripheral Nerve Injuries Producing Loss of Function. Brain *74*, 491–516.

Švec, J.G., and Schutte, H.K. (1996). Videokymography: High-speed line scanning of vocal fold vibration. Journal of Voice *10*, 201–205.

Takahara, T., and Kwee, T.C. (2012). Low b-value diffusion-weighted imaging: Emerging applications in the body. Journal of Magnetic Resonance Imaging *35*, 1266–1273.

Taoka, T., Sakamoto, M., Nakagawa, H., Fukusumi, A., Iwasaki, S., Taoka, K., and Kichikawa, K. (2005). Evaluation of Extraocular Muscles Using Dynamic Contrast Enhanced MRI in Patients With Chronic Thyroid Orbitopathy. Journal of Computer Assisted Tomography January *29*, 115–120.

Temple, L., Fuzesi, S., and Patil, S. (2009). The importance of determining quality of life in clinical trials. Surgery *145*, 622–626.

Thoma, A., Farrokhyar, F., McKnight, L., and Bhandari, M. (2010). How to optimize patient recruitment. Can J Surg *53*, 205–210.

Tian, W., Yin, H., Redett, R.J., Shi, B., Shi, J., Zhang, R., and Zheng, Q. (2010). Magnetic resonance imaging assessment of the velopharyngeal mechanism at rest and during speech in Chinese adults and children. J. Speech Lang. Hear. Res. *53*, 1595–1615.

Titze, I. (1995). Workshop on Acoustic Voice Analysis; Summary Statement. National Centre for Voice and Speech, The University of Iowa, IA.

Toran, K., and Lal, B. (2010). Objective analysis of voice in normal young adults. Kathmandu University Medical Journal *7*, 374–377.

Tucker, H.M. (1976). Human Laryngeal Reinnervation. The Laryngoscope *86*, 769–779.

Tucker, H.M. (1989). Long-term results of nerve-muscle pedicle reinnervation for laryngeal paralysis. Ann. Otol. Rhinol. Laryngol *98*, 674–676.

Tucker, H.M., and Rusnov, M. (1981). Laryngeal reinnervation for unilateral vocal cord paralysis: long-term results. Ann. Otol. Rhinol. Laryngol *90*, 457–459.

Uloza, V., Pribuisiene, R., and Saferis, V. (2005). Multidimensional assessment of functional outcomes of medialization thyroplasty. Eur Arch Otorhinolaryngol *262*, 616–621.

Uloza, V., Vegienė, A., Pribuišienė, R., and Šaferis, V. (2013). Quantitative Evaluation of Video Laryngostroboscopy: Reliability of the Basic Parameters. Journal of Voice *27*, 361–368.

Urquhart, A.C., Johnson, J.T., Myers, E.N., and Schechter, G.L. (1994). Glomus Vagale. The Laryngoscope *104*, 440–445.

Vachha, B., Cunnane, M.B., Mallur, P., and Moonis, G. (2013). Losing your voice: etiologies and imaging features of vocal fold paralysis. J Clin Imaging Sci *3*, 15.

Verikas, A., Uloza, V., Bacauskiene, M., Gelzinis, A., and Kelertas, E. (2009). Advances in laryngeal imaging. European Archives of Oto-Rhino-Laryngology *266*, 1509–1520.

Verin, E., Marie, J.-P., and Similowski, T. (2011). Cartography of human diaphragmatic innervation: Preliminary data. Respiratory Physiology & Neurobiology *176*, 68–71.

Verstraete, K.L., Woude, H.-J.V. der, Hogendoorn, P.C.W., De Deene, Y., Kunnen, M., and Bloem, J.L. (1996). Dynamic contrast-enhanced MR imaging of musculoskeletal tumors: Basic principles and clinical applications. Journal of Magnetic Resonance Imaging *6*, 311–321.

Viddeleer, A.R., Sijens, P.E., Ooyen, P.M.A. van, Kuypers, P.D.L., Hovius, S.E.R., and Oudkerk, M. (2012). Sequential MR Imaging of Denervated and Reinnervated Skeletal Muscle as Correlated to Functional Outcome. Radiology *264*, 522–530.

Viera, A.J., and Garrett, J.M. (2005). Understanding Interobserver Agreement: The Kappa Statistic. Fam Med *37*, 360–363.

Vinson, K.N., Zraick, R.I., and Ragland, F.J. (2010). Injection versus medialization laryngoplasty for the treatment of unilateral vocal fold paralysis. The Laryngoscope *120*, 1802–1807.

Volk, G.F., Hagen, R., Pototschnig, C., Friedrich, G., Nawka, T., Arens, C., Mueller, A., Foerster, G., Finkensieper, M., Lang-Roth, R., et al. (2012). Laryngeal electromyography: a proposal for guidelines of the European Laryngological Society. Eur Arch Otorhinolaryngol *269*, 2227–2245.

Wang, C.-T., Liao, L.-J., Huang, T.-W., and Cheng, P.-W. (2012). Preliminary Report of Vocal Fold Augmentation with Cross-Linked Porcine Collagen. Otolaryngology -- Head and Neck Surgery *146*, 606–610.

Wang, W., Chen, D., Chen, S., Li, D., Li, M., Xia, S., and Zheng, H. (2011). Laryngeal reinnervation using ansa cervicalis for thyroid surgery-related unilateral vocal fold paralysis: a long-term outcome analysis of 237 cases. PLoS ONE *6*, e19128.

Waters, K.A., Woo, P., Mortelliti, A.J., and Colton, R. (1996). Assessment of the infant airway with videorecorded flexible laryngoscopy and the objective analysis of vocal fold abduction. Otolaryngology - Head and Neck Surgery *114*, 554–561.

Webb, A.L., Carding, P.N., Deary, I.J., MacKenzie, K., Steen, N., and Wilson, J.A. (2003). The reliability of three perceptual evaluation scales for dysphonia. Eur Arch Otorhinolaryngol *261*.

Weber, M.-A., Krix, M., and Delorme, S. (2007). Quantitative evaluation of muscle perfusion with CEUS and with MR. Eur Radiol *17*, 2663–2674.

Wen, M.-H., Cheng, P.-W., Liao, L.-J., Chou, H.-W., and Wang, C.-T. (2013). Treatment Outcomes of Injection Laryngoplasty Using Cross-Linked Porcine

Collagen and Hyaluronic Acid. Otolaryngology -- Head and Neck Surgery *149*, 900–906.

Wessig, C., Koltzenburg, M., Reiners, K., Solymosi, L., and Bendszus, M. (2004). Muscle magnetic resonance imaging of denervation and reinnervation: correlation with electrophysiology and histology. Experimental Neurology *185*, 254–261.

Woodson, G.E. (1993). Configuration of the glottis in laryngeal paralysis. I: Clinical study. The Laryngoscope *103*, 1227–1234.

Woodson, G.E. (2007). Spontaneous laryngeal reinnervation after recurrent laryngeal or vagus nerve injury. Ann. Otol. Rhinol. Laryngol *116*, 57–65.

Yamabe, E., Nakamura, T., Oshio, K., Kikuchi, Y., Ikegami, H., and Toyama, Y. (2008). Peripheral nerve injury: diagnosis with MR imaging of denervated skeletal muscle--experimental study in rats. Radiology *247*, 409–417.

Yanagisawa, O., Shimao, D., Maruyama, K., Nielsen, M., Irie, T., and Niitsu, M. (2009a). Diffusion-weighted magnetic resonance imaging of human skeletal muscles: gender-, age- and muscle-related differences in apparent diffusion coefficient. Magnetic Resonance Imaging *27*, 69–78.

Yanagisawa, O., Shimao, D., Maruyama, K., and Nielsen, M. (2009b). Evaluation of exercised or cooled skeletal muscle on the basis of diffusion-weighted magnetic resonance imaging. Eur J Appl Physiol *105*, 723–729.

Yiu, E.M.L., Lau, V.C.Y., Ma, Estella, P.-M., Chan, K.M.-K., and Barrett, E. (2013). Reliability of laryngo-stroboscopic evaluation on lesion size and glottal configuration: A revisit. The Laryngoscope *124*, 1638–1644.

Yılmaz T, Süslü N, Atay G, Özer S, Günaydın R, and Bajin M (2013). Comparison of voice and swallowing parameters after endoscopic total and partial arytenoidectomy for bilateral abductor vocal fold paralysis: A randomized trial. JAMA Otolaryngol Head Neck Surg *139*, 712–718.

Young, J.M., Solomon, M.J., Harrison, J.D., Salkeld, G., and Butow, P. (2008). Measuring patient preference and surgeon choice. Surgery *143*, 582–588.

Yumoto, E., Minoda, R., Hyodo, M., and Yamagata, T. (2002). Causes of recurrent laryngeal nerve paralysis. Auris Nasus Larynx *29*, 41–45.

Yumoto, E., Sanuki, T., Toya, Y., Kodama, N., and Kumai, Y. (2010). Nerve-muscle pedicle flap implantation combined with arytenoid adduction. Arch. Otolaryngol. Head Neck Surg *136*, 965–969.

Zealear, D.L., Hamdan, A.L., and Rainey, C.L. (1994). Effects of denervation on posterior cricoarytenoid muscle physiology and histochemistry. Ann. Otol. Rhinol. Laryngol. *103*, 780–788.

Zhang, Y., and Jiang, J.J. (2008). Acoustic Analyses of Sustained and Running Voices From Patients With Laryngeal Pathologies. Journal of Voice *22*, 1–9.

Zhao, S.-X., Tsui, S., Cheung, A., Douglas, R.S., Smith, T.J., and Banga, J.P. (2011). Orbital fibrosis in a mouse model of Graves' disease induced by genetic immunization of thyrotropin receptor cDNA. J Endocrinol *210*, 369–377.

Zheng, H., Li, Z., Zhou, S., Cuan, Y., and Wen, W. (1996). Update: laryngeal reinnervation for unilateral vocal cord paralysis with the ansa cervicalis. Laryngoscope *106*, 1522–1527.

Zur, K.B. (2012). Recurrent laryngeal nerve reinnervation for unilateral vocal fold immobility in children. Laryngoscope *122 Suppl 4*, S82–S83.

Appendix 1- ENT UK survey questionnaire

Laryngeal reinnervation versus thyroplasty in unilateral vocal fold paralysis

1. Do you see patients with unilateral vocal fold paralysis in your clinical practice?

Yes/ No

2. What proportion of your patients are adult laryngology patients?

Less than 10%	
10 to 50%	
More than 50%	

3. How many cases a year do you see conforming to the following descriptions?

An adult patient with unilateral vocal fold paralysis (UVFP)	
Adult UVFP showing no voice improvement following voice therapy	

4. What treatment option(s) would you consider for an adult patient with unilateral vocal fold paralysis, showing no voice improvement following voice therapy?

Injection laryngoplasty	
Isshiki Type I Thyroplasty	
Laryngeal Reinnervation with or without medialisation	
Refer to other ENT Consultant/ Laryngologist (please specify the names and hospital of colleagues to whom you would refer in the box	
Other (please specify)	

5. Please indicate your preferred type of anaesthesia for the surgical options.

	Preferred type of anaesthesia	
	General anaesthesia	Local anaesthesia
Injection laryngoplasty		
Isshiki type I thyroplasty		

6. If you do Isshiki Type I Thyroplasty, what would be your choice of implant?

Silastic block implant	
Montgomery implant	
Gore-tex implant	
Titanium implant	
Others(please specify)	

7. Have you performed any laryngeal reinnervation surgery in the last 5 years?

Yes	

8. If a trial comparing the effectiveness of laryngeal reinnervation versus Isshiki type I thyroplasty for adult patients with UVFP due to isolated permanent recurrent laryngeal nerve paralysis showing no voice improvement following voice therapy was initiated, would you be willing to enter patients into such a trial?

No. Please give reason in the box below Please provide reason if your answer is 'No'	

9. If you would like to participate in this trial, please provide your details below.

Name	
Hospital	
Place	

10. If your patients are willing to enter the trial, would you

Refer my patient to the principal investigator for both types of surgery	
Perform Isshiki type I thyroplasty in my own centre, but refer my patient to the principal investigator for laryngeal reinnervation	
Perform both sorts of surgery in my centre	

11. Would you like to receive training in

	Yes	No
Isshiki Type I Thyroplasty		
Laryngeal reinnervation		

12. What kind of hospitals are you practising in?

District General Hospital	
University Teaching Hospital	
Private Hospital	

13. Which of these tests are available in your centre?

Videolaryngostroboscopy	
Acoustic studies	
Aerodynamic studies	
Laryngeal electromyography	
Fiberoptic endoscopic examination of swallowing (FEES)	
Videoflouroscopic swallow evaluation	
Magnetic resonance imaging	

14. Please feel free to make any further comments below

Appendix 2 – Participant information sheet

(for patients with voice box weakness**)**

Study title : Laryngeal Reinnervation versus Type I Thyroplasty (**PhD Research Project**)

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it would involve for you. **One of our team will go through the information sheet with you and answer any questions you have.** We suggest this should take about 10 minutes.

Before you decide whether or not to take part, it is important to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. **Ask us if there is anything that is not clear**, or if you would like more information.

(Part 1 tells you the purpose of this study and what will happen to you if you take part.
Part 2 gives you more detailed information about the conduct of the study).

Thank you for reading this.

PART 1

1.1 What is the purpose of this study?

One of the nerves to your voice box (larynx) is not working well and this is affecting your voice and possibly your swallowing and breathing. We want to compare two different operations used to treat this problem. In the first operation, we repair the neck nerves using parts of other, unaffected nerves instead (‘laryngeal reinnervation’). The alternative (traditional) method is to move around bits of tissue (cartilage) to give you a better voice. We do not know which of these methods is better and so invite you consider taking part in our research.

1.2 Why have I been invited?

You have been invited because you have voice box weakness and may be swallowing problems due to nerve injury or of unknown cause. We need to do one of the above procedures to improve these problems. There will be 12 participants in total for this study

1.3 Do I Have To Take Part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or not to take part, will not affect the standard of care you receive.

1.4 What Will Happen To Me If I Take Part?

Sometimes we don't know which way of treating patients is best. To find out, we need to compare different treatments. We put people into groups and give each group a different treatment. The results are compared to see if one is better. To try to make sure the groups are the same to start with, each participant is put into a group by chance (randomly, so this is called 'randomisation'). You have a 50-50 chance undergoing either treatment or operation. The two different operations are explained below:

One operation ('thyroplasty') is carried out under local anaesthesia, with some medicine that makes you sleepy and relaxed. A small cut will be made on your neck. A small block made from a safe type of plastic will be placed into the side of your voice box, and this pushes the weak vocal cord into a position which is better for speech and swallowing. You would not have pain during the operation and you would not remember the operation. The operation lasts about 1 hour.

The other operation ('reinnervation') is to improve the strength of the muscles in the voice-box and is carried out under general anaesthesia. A small cut will be made on your neck. During the operation, the nerve supplying your voice box will identified and will be connected to the nerve which supplies one of the muscles on the front of your neck. Any changes in the way this neck muscle works after the operation are

not noticeable. At the same time a safe substance made of collagen will be injected into the weak vocal cord using a telescope through your mouth. The collagen injection is not permanent, but will give you some useful voice improvement while we wait for the result of the nerve rewiring to take effect. The operation lasts about 2 hours.

After either operation, you will be kept overnight at the RNTNE Hospital. Then, we will see you in clinic 3 month, 6 months and 12 months afterwards so that we can see how you are getting on.

Before and after the operation we will:

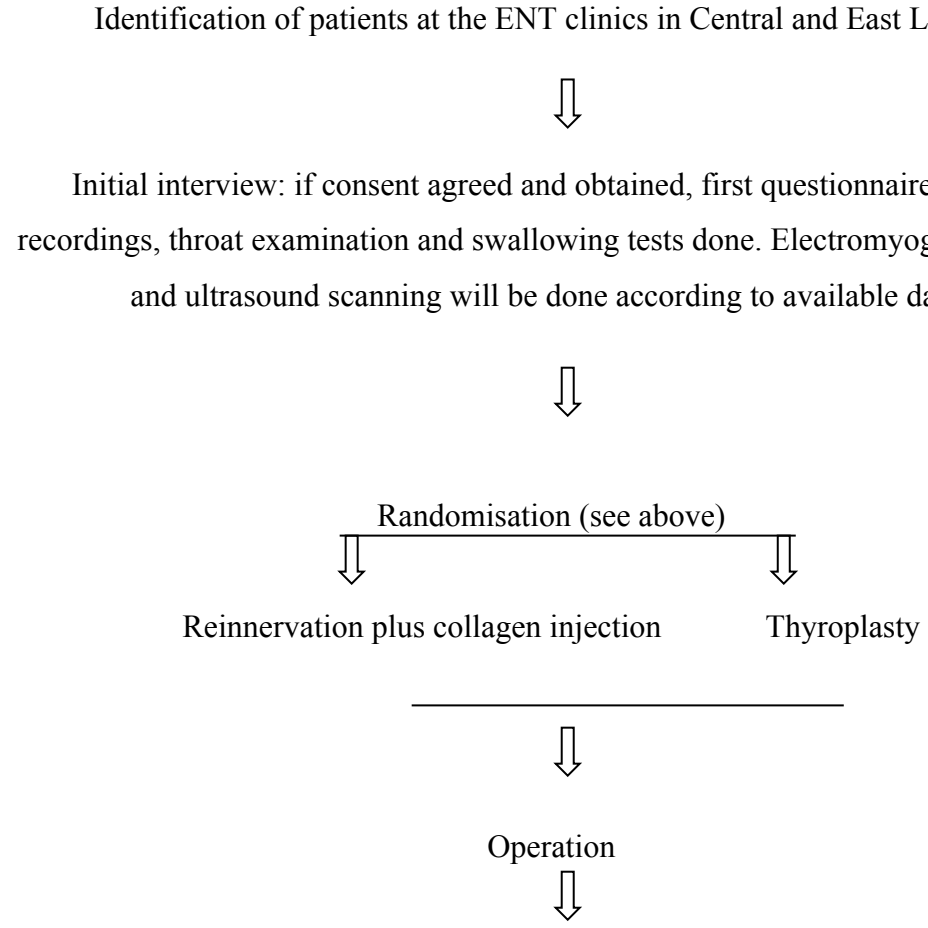
- record your voice at speech room as well using and a specially designed easy-to-use portable voice recorder on an Apple ipod touch - OperaVOX™
- record your vocal cord movements using a camera connected to a telescope passed through the nose
- measure your swallowing
- measure how well the muscles are recovering using a special study described below (electromyography)
- measure your breathing function by flow volume loop test

Electromyography will be done in the clinic with you lying down. It will be carried out by experienced doctors. One doctor will put tiny needles into the voice-box muscles being tested through the front part of neck. You will be asked to make sounds to test how each muscle moves. Typically, this takes about 10 minutes in total. You may experience brief pain or discomfort during insertion of the needles (similar to that when you have a blood test at the doctors), but most people find the tests fine. Rarely, there is temporary mild bleeding or bruising at the needle sites.

We wish to also do an extra test that is not done normally for patients with your problem. The tests are called magnetic resonance imaging (MRI) that will be carried out at the University College London Hospital. For MRI, you lie down in a dark tube for up to 1 hour whilst the machine takes pictures of the voice box. These are like X-

rays, but do not have the same risks as X-rays and give better pictures. However, some people find that lying in the tube is an uncomfortable experience. It is also noisy, but you will be given headphones to make it quieter and music can be played through these. MRI will be done before surgery and at 12 months after surgery.

Diagram of events in this study



Clinic visits at 3 month, 6 months and 12 months at the RNTNEH for more questionnaires, voice recordings and throat examination. Electromyography and MRI will be repeated at 12months after operation. Voice recordings using OperaVOX™ will be done at home monthly.

1.5 Expenses and payments

We will reimburse the travel expenses during visits either to RNTNEH or University College London Hospital for the purpose of this study.

1.6 What will I have to do?

People who have thyroplasty operations will be given an antibiotic called erythromycin. Please inform us if you are allergic to it for us to substitute it with other type of antibiotic. It is best if you stick to the dates of appointment for clinics. This is important for us to compare the changes after surgery between different people and operations. You need to let us know if there are any changes in your voice and swallowing problems after the surgery during the clinic visits. You will be asked to fill in questionnaires about these changes.

It is important to inform the doctor if you are on any blood-thinning medications or if you have a bleeding disorder prior to having the electromyography performed. You are encouraged to wear a shirt or top that allows access to the lower neck and to avoid neck jewellery on the date of the exam.

1.7 What are the alternatives for treatment?

The alternative treatment available is injection of some sort of filling material into the affected vocal fold using a special instrument through mouth. However this treatment measure is not long lasting in most people and the injection may need to be repeated every few months.

1.8 What are the side effects or disadvantages of taking part?

Reinnervation and thyroplasty are established operations to treat a paralysed voice box (larynx) but there is not enough information to tell us which one is better since a study like this one has not been published yet.

We do not predict any significant disadvantages in taking part in this study. If you are selected for the first operation, it will be done under general anaesthesia instead of local anaesthesia. An anaesthetist will do an assessment to identify any higher risks than normal of your having anaesthetic side effects.

Both are safe operations with no significant risk or side effects reported in the literature. Reinnervation has a delayed effect because the rewired nerve needs at least three months to regain its function. However the delay in its effect should not be

noticeable as we will also inject some fat to temporarily improve the position of the vocal cord as described.

Possible side effects of the operations are very similar: swelling of the neck, blood clot accumulation or infection of the wound, and difficulty in breathing may occur. These may occur with any form of surgery in the neck area, and are not uncommon. The chances of these happening are the same whichever operation you receive.

You need to undergo two extra tests as part of this study which you would not have received if you were receiving treatment outside a trial. The tests are Magnetic Resonance Imaging (MRI) and ultrasound scanning which are described above. They will be performed at the University College London Hospital on the same day. We will repay any costs you have to pay out to attend these scans.

With regards to the MRI and ultrasound scanning, there would be no risk of exposure to radiation. However some people might experience claustrophobia (a feeling of stuffiness due to being in a small space) with MRI.

1.9 What are the possible benefits of taking part?

The main benefit is the knowledge that you are helping to improve care for future patients who may have the problem and who need the operation. By taking part in this study we will know which procedure is more effective in improving the voice and swallowing.

1.10 What happens when the research study stops?

You will be getting regular clinic visits after the trial to monitor the long term changes. Additional treatment will be provided if necessary. The monitoring will be done by the chief investigator and his team.

You will be informed about the continuation of treatment after the trial. This would not incur any additional costs since it is part of daily clinical practice. The conclusion of the trial will be presented at an international conference and will be sent for publication. The conclusion of this study will indicate the feasibility of a bigger trial which will be carried out accordingly and in due course. The trial results

will provide some evidence of efficacy because randomisation will be performed but it will be inadequate due to the small number of subjects.

1.11 What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

1.12 Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

PART 2

2.1 What if new information becomes available?

Sometimes during the course of a research project, new information becomes available about the treatment that is being studied. If this happens, the study doctor will tell you about it and discuss with you whether you want to continue to take part. If you decide to withdraw, your operation will proceed as normal. If you decide to continue, you will be asked to sign an updated consent form.

2.2 What will happen if I don't want to carry on with the study?

At any stage prior to surgery, you can decide to withdraw from the study. You can withdraw from treatment but keep in contact with us to let us know your progress. Information collected may still be used. Your doctor will look after you normally, the same as if the study were not happening at all. Withdrawing from the study will not affect your treatment in any way and you will not offend anybody by doing so. You are also free to ask the opinion of your GP, or other doctors and nurses about the study if you wish.

2.3 What if there is a problem?

Every care will be taken in the course of this study. However, in the unlikely event that you are injured by taking part, compensation may be available.

If you suspect that the injury is the result of the Sponsor's (University College London) or the hospital's negligence then you may be able to claim compensation. After discussing with your research doctor, please make the claim in writing to the

Professor Martin Birchall who is the Chief Investigator for the research and is based at the RNTNEH. The Chief Investigator will then pass the claim to the Sponsor's Insurers, via the Sponsor's office. You may have to bear the costs of the legal action initially, and you should consult a lawyer about this.

Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated by members of staff or about any side effects (adverse events) you may have experienced due to your participation in the research, the normal National Health Service complaints mechanisms are available to you. Please ask your research doctor if you would like more information on this. Details can also be obtained from the Department of Health website: http://www.dh.gov.uk.

2.4 Will my taking part in this study be kept confidential?

Information collected about you during the course of the research will be kept strictly confidential. Any data that have your identifiable personal data will be removed. Names would not be written in full but with the first letter of given name and surname with number coding. Electronic data will be kept in computers secured with password and non-electronic data will be kept in a locked cabinet. The data is only accessible by the research team members. No data will be kept in laptops. If there is pressing need to do so, the data will be saved in an encrypted file. The personal data confidentiality will follow the NHS Code of Confidentiality.

In accordance with current UCL Records Management Policy, research findings will need to be stored by UCL as sponsor for 20 years after the research has finished. The UCL Records Office provides a service to UCL staff and maintains archived records in a safe and secure off site location. All activities are conducted in accordance with the Data Protection Act and UCL Data Protection Policy. Access to the data is strongly regulated and permissions to access the data are treated case by case.

2.5 Involvement of the General Practitioner/Family doctor (GP)

Your GP will be notified of your participation in this study with your permission.

2.6 What will happen to any samples I give?

There wouldn't be any blood or tissue sample taken. Your voice and voice box findings will be recorded to monitor and analyse the changes. The recordings will be kept in computers encrypted with password.

2.7 What will happen to the results of the research study?

We hope to recruit twelve participants into this study. Once it has all been completed we will present the results at an international meeting attended by doctors from around the world. After this, we will publish the results in an appropriate scientific journal. This could take about a year in total. If you wish, we will make sure you receive your own copy of the results and paper as soon as we have them.

2.8 Who is organising and funding the research?

The sponsor is the University College London. So far the research is funded by the internal funding. The internal funder will pay RNTNEH for including you in this study.

2.9 Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Research Ethics Committee.

You will be given a copy and a signed consent form to keep.

2.10 What if I have other concerns?

The general information about this research will be available at the Ear Institute website http://www.ucl.ac.uk/silva/ear/staff/marinamatbaki. To get the specific information about this research project, you need to contact the researcher (Professor Birchall or Marina - 020 7915 1308). Alternatively, you may email as below.

If you have any concerns or questions, you may contact any one of the researchers listed below for advice at the Royal National Throat Nose and Ear Hospital, tel: 020 7915 1308. Leave your contact number to Anne Oliphant for the researchers to return your call.

Professor Martin Birchall, Consultant and Professor of Head and Neck Surgery (m.birchall@ucl.ac.uk)

Dr Marina Mat Baki, PhD Student in laryngology (sjjgmbm@live.ucl.ac.uk)

Appendix 3 – Voice Handicap Index (VHI-10)

Instructions: These are statements that many people have used to describe their voices and the effects of
their voices on their lives. Circle the response that indicates how frequently you have the same experience.

0 = never 1 = almost never 2 = sometimes 3 = almost always 4 = always

Fl. My voice makes it difficult for people to hear me.	0	1	2	3	4
F2. People have difficulty understanding me in a noisy room	0	1	2	3	4
F8. My voice difficulties restrict personal and social life	0	1	2	3	4
F9. I feel left out of conversations because of my voice	0	1	2	3	4
F10. My voice problem causes me to lose income	0	1	2	3	4
P5. I feel as though I have to strain to produce voice	0	1	2	3	4
P6. The clarity of my voice is unpredictable	0	1	2	3	4
E4. My voice problem upsets me	0	1	2	3	4
E6. My voice makes me feel handicapped	0	1	2	3	4
P3. People ask, “What’s wrong with your voice?”	0	1	2	3	4

VHI-10 = _____

Appendix 4 - Eating Assessment Tool (Eat-10)

Tick the appropriate response
To what extent are the following scenarios problematic for you?

		0 = no problem 4 = severe problem				
1	My swallowing problem has caused me to lose weight	0	1	2	3	4
2	My swallowing problem interferes with my ability to go out for meals	0	1	2	3	4
3	Swallowing liquids takes extra effort	0	1	2	3	4
4	Swallowing solids takes extra effort	0	1	2	3	4
5	Swallowing pills takes extra effort	0	1	2	3	4
6	Swallowing is painful	0	1	2	3	4
7	The pleasure of eating is affected by my swallowing	0	1	2	3	4
8	When I swallow food sticks in my throat	0	1	2	3	4
9	I cough when I eat	0	1	2	3	4
10	Swallowing is stressful	0	1	2	3	4

Appendix 5- Stroboscopy Research Instrument

Stroboscopy Research Instrument

Normal = 0; Mild = 1; Moderate = 2; Severe = 3

Symmetry of mucosal displacement

	normal	mild	mod	severe
Asymmetry				

Is the horizontal excursion away from midline reduced?

Amplitude	normal	mild	mod	severe	none
Right					
Left					

How irregular is the frequency of vibration of each vocal fold?

Periodicity	normal	mild	mod	severe	none
Right					
Left					

Mucosal segments that do not participate in vibratory activity

Non-vibratory segment	none	Anterior 1/3	Middle 1/3	Posterior 1/3	The whole vocal
Right					
Left					

Duration of mucosal contact during vocal fold vibration

Duration of closure	Predominately closed	½ closed ½ opened	Predominately opened	Always open

Closure
Predominant mucosal closure pattern

Closure *Choose one closure*	Present	*If present ,please choose:* Mild	Moderate	Severe
Hourglass 				
Spindle 				
Posterior Glottic Chink 				
Anterior Glottic Chink 				
Complete Closure 				
Complete Nonclosure 				

_____ Check if you viewed this case in slow-motion.

Appendix 6 – Stage I and II MRI study

Stage I: Protocol development

5 healthy volunteers who do not have any voice problems were invited to participate. A protocol for anatomical metrics (T2 weighted imaging, T1 weighted imaging) and functional metrics (diffusion weighted imaging, dynamic contrast enhanced imaging, cine imaging for vocal fold mobility) on 3T MRI will be determined. The final protocol was then used for Phase II, feasibility study to find out whether it is feasible on UVFP patients. The final protocol that comprises T1-weighted imaging (non-fat suppression), post-contrast T1-weighted imaging (fat suppression), T2-weighted imaging, diffusion weighted imaging (DWI), dynamic contrast enhanced MRI (DCE-MRI) and cine-MRI (Table 1)

Table 1: Protocol developed for MRI larynx on 3Tesla machine

	T1 weighted	T2 weighted	DWI	DCE-MRI	Cine-MRI
Sequence Name	SE	SE-TSE	SE-EPI	FFE	FFE
Repetition time (ms)	450	4000	2000	4.4	1.98
Echo time (ms)	10	100	80	2.2	0.78
Image resolution	348x310	300x297	116x115	188x186	160x120
Slice thickness	2.5mm	2.5mm	5mm	3mm	10mm
Number of slices	25	28	14	24	3
No of averages	6	4	6	1	1
Diffusion weighting (b-value)	n/a	n/a	0, 50, 100, 300, 600, 1000	n/a	n/a
Temporal resolution (s)	n/a	n/a	n/a	19.1	0.736
Total acquisition time (mins)	9min 9sec	10min 42sec	10min 4sec	11min 10sec	1.28

Stage II: Feasibility study

Five healthy volunteers and 5 UVFP patients identified from ENT clinics were invited to participate. These patients were scanned using a 3T MRI scanner using the final protocol in table 4 that has been developed in stage I.

Imaging was performed on a 3T Philips Achieva multi-transmit MRI scanner using the manufacturers 16 channel head and neck coil. MRI was performed to evaluate macrostructure (muscle bulk and fatty change – T1/T2 weighted imaging), microstructure (cellularity/perfusion – diffusion weighted imaging), vascularity (perfusion characteristics – dynamic contrast enhanced imaging) and motion (movement during phonation – cine MRI) as per protocol from stage I study. All of the patients had laryngeal electromyography performed to confirm vocal cord paralysis and the findings were classified according to Koufman and Walkers classification (Koufman et al., 2001).

RESULTS

A. Macrostructure

I. Image quality assessment

Two head and neck radiologist consultants review the images of 5 healthy volunteers and 4 UVFP patients acquired using the protocol developed. Majority of the coronal T2-weighted images (Figure 1, Figure 2) and cine-MRI (Figure 3) for vocal cord motion had been scored as good and for the axial T2-weighted images majority was excellent. None of the images were rated as very poor. On the other hand, majority of the T1 weighted images (Figure 4, Figure 5) were rated as fair with some of them were of very poor quality.

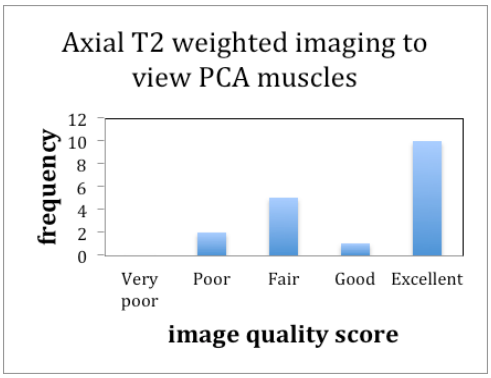


Figure 1: Image quality evaluation of coronal T2-weighted imaging for TA muscle

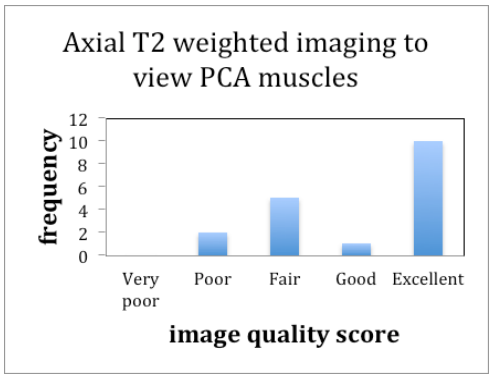


Figure 2: Image quality evaluation of axial T2-weighted imaging for PCA muscle

Figure 3: Image quality evaluation of cine imaging for vocal fold mobility

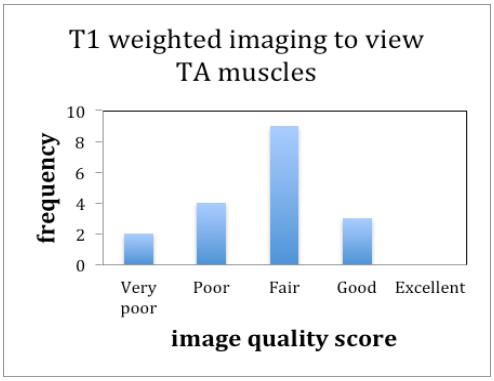


Figure 4: Image quality evaluation of T1-weighted imaging for TA muscle

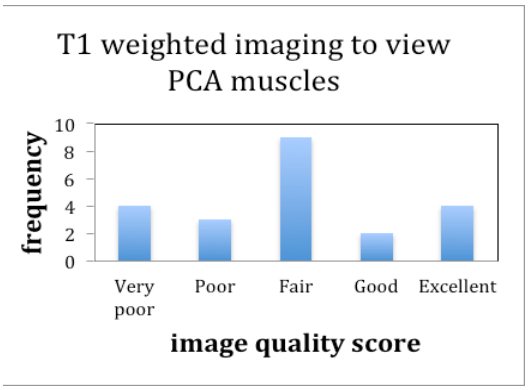


Figure 5: Image quality evaluation of T1-weighted imaging for PCA muscle

II. Quantitative result

Five UVFP patients and 5 healthy volunteers with matched age had undergone MRI scan on a 3Tesla MRI machine according to the protocol developed. Only 3 patients had their TA muscles analysed because the other 2 patients had injection laryngoplasty to their vocal folds. T2-weighted imaging generally showed muscle of the paralysed PCA was smaller than the opposite normal and the volunteers' PCA. The absolute value of difference (ABSD) between paralysed and normal muscle in patients was significantly larger compared to volunteers (Table 2) except in one patient. There were also abnormal signal changes on the paralysed muscles (Figure 5). The signal was more hyperintense on the paralysed muscles therefore the ratio of mean signal intensity of normal muscles to the opposite paralysed muscle was significantly smaller than the volunteers' right to left normal muscles (Table 2). Similar findings were also seen on TA muscles (Figure 7)(Table 3).

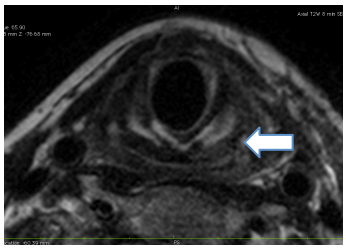


Figure 6: T2 weighted axial images through the vocal cord muscles in a patient with left vocal cord paralysis (arrowhead). Note increased relative T2 signal of the paralysed left PCA muscle

Figure 7: T2 weighted coronal images through the vocal cord muscles in a patient with left vocal cord paralysis (arrowhead). Note increased relative T2 signal and smaller bulk of the paralysed left TA muscle

Table 2: Muscle bulk and signal intensity of posterior cricoarytenoid muscle.

Variables	Vocal fold paralysis (5 subjects, 5 paralysed vocal folds)	Volunteers (5 subjects, 10 normal vocal folds)
Age, mean (SD)	44.0 (16.53)	32.4 (5.98)
Independent t-test	P = 0.187	
Gender	Female (4)	Female (3)
Muscle bulk		
Median (IQR)	0.19 (0.16, 0.34)	0.493 (0.3, 0.5)
Mann-Whitney test	P = 0.008	
Median (IQR) ABSD	0.97 (0.43, 01.67)	0.28 (0.005, 0.09)
Mann-Whitney U test	P = 0.009	
Signal intensity Ratio R vs L or N vs P		
Median (IQR)	0.61 (0.55, 0.82)	1.01(0.97, 1.19)
Mann-Whitney U test	P = 0.016	

ABSD: Absolute value of difference
SD: Standard deviation
IQR: interquarile range

Table 3: Muscle bulk and signal intensity of thyroarytenoid muscle.

Variables	Vocal fold paralysis (5 subjects, 5 paralysed vocal folds, 2 missing data)	Volunteers (5 subjects, 10 normal vocal folds)
Muscle bulk		
Median (IQR)	0.23 (0.21, 0.27)	0.40 (0.39, 0.80)
Mann-Whitney U test	P = 0.01	
Median (IQR) ABSD	0.72 (0.63, 01.79)	0.04 (0.01, 0.08)
Mann-Whitney U test	P = 0.025	
Signal intensity Ratio R vs L or N vs P		
Median (IQR)	0.47 (0.46, 0.51)	1.08 (1.0, 1.15)
Mann-Whitney U test	P = 0.025	

ABSD: Absolute value of difference
SD: Standard deviation
IQR: interquarile range
R : right, L : left, N : normal, P : paralysed

B. Microstructure

I. Diffusion weighted imaging (DWI)

Mean apparent diffusion coefficient (ADC) value for low b values that represent perfusion of healthy volunteers was 2.4±0.5 $x10^{-3}mm^2s^{-1}$ (Table 4). The ADC value of paralysed vocal fold increased in one patient whereas in the other 2 patients, it was decreased. This preliminary result indicates a higher perfusion in 1 patient whereas the other 2 patients showed lower perfusion related ADC on the paralysed side of the TA muscles. Reliable signal of PCA muscles could not be acquired on the DWI. Future work will focus on low b values of DWI and improvement of signal to noise ratio.

Table 4: ADC thyroarytenoid muscles of volunteers

Thyroarytenoid muscle Volunteer		ADC of low b values (0, 50, 100) $x10^{-3}mm^2s^{-1}$	
		Right	Left
	V01	2.15	2.13
	V02	2.35	2.52
	V03	2.19	1.60
	V04	1.88	2.80
	V05	3.12	3.17
Mean (SD)		2.4 (0.5)	

Table 5: ADC thyroarytenoid muscles of patients

Thyroarytenoid muscle Patient		ADC 0f low b values (0, 50, 100) $x10^{-3}mm^2s^{-1}$	
		Right	Left
	P01	4.22	6.71
	P02	injection	injection
	P03	injection	injection
	P04	2.28	1.47
	P05	2.86	2.04

Injection : injection laryngoplasty done on the vocal cords

II. Dynamic contrast enhancement imaging (DCE)

Time intensity curve (TIC) was generated for all participants except one patient who had too much movement on the TA muscle during image acquisition that a reasonable TIC could not be generated. Differences of DCE parameters between two vocal fold muscles of the same subject were shown in ratio 'right to left' in volunteers and 'normal to paralysed' vocal cord in patients (Table 6, Table 7). A bigger ratio 'more than 1' in patients indicates a bigger value in the normal than the paralysed vocal folds. Mean ratio 'close to one' indicates that there is not much difference between the 2 vocal cords. Mean ratio of slope, ME and AUC in volunteers between right and left of PCA and TA muscles were closed to 1. Mean ratio of slope in patients was 1.67, far 'larger than 1', which indicates lower rate of initial enhancement of the paralysed muscles than the opposite normal muscles although it was not statistically significant different from the volunteers group. The mean of ME and AUC of patients were close to one. However if we analyse the parameters by case-to-case basis (Table 8), 3 patients with high vagal paralysis showed lower slope, ME and AUC of the PCA muscles than the normal opposite PCA muscles whereas the other 2 patients who had recurrent laryngeal nerve paralysis had almost similar values of all the metrics except the slope in patient 05. On the other hand, these 2 patients' paralysed TA muscles showed smaller slope and AUC than the opposite normal TA muscles. Regarding patterns of TIC of both normal TA and PCA muscles of healthy volunteers after reaching the initial maximum, most of them showed type B, 3 had type A and 1 had type C. Normal vocal fold muscles in patients showed mostly type B except in 2 patients, 1 was type A and the other one was type C. The opposite paralysed vocal folds showed the similar type of TIC but with smaller slope, ME or AUC. All of the TIC patterns of

the paralysed vocal cord muscles did not show evidence of wash-out pattern except in the paralysed PCA of patient 04.

Table 6: Ratio of DCE parameters of posterior cricoarytenoid (PCA) muscles

Mean (SD) ratio of DCE parameters – PCA muscles			
	SoE	ME	AUC
Volunteers			
Mean (SD)	1.03 (0.17)	1.07 (0.08)	0.95 (0.07)
Median (IQR)	0.99 (0.74,1.35)	1.02 (0.94,1.23)	0.94 (0.90,1.01)
Patients			
Mean (SD)	1.67 (0.71)	1.14 (0.23)	1.07 (0.11)
Median (IQR)	1.51 (1.16,2.26)	1.05 (0.94,1.38)	1.11 (0.94,1.17)
Mann Whitney U test	p = 0.175	p = 0.754	p = 0.117

SD: Standard deviation
IQR: interquartile range

Table 7: Ratio of DCE parameters of thyroarytenoid (TA) muscles

mean (SD) ratio of DCE parameters – TA muscles			
	slope	ME	AUC
Volunteers	1.11 (0.33)	1.04 (0.13)	1.01 (0.25)
Patient 04	1.92	0.73	1.14
Patient 05	1.71	1.17	1.14

SD: Standard deviation
DCE: Dynamic contrast enhancement
ME: Maximum enhancement AUC:
Area under the curve

Table 8: Summary of UVFP patients results

	Patient 01	Patient 02	Patient 03	Patient 04	Patient 05
Age	40	18	52	61	51
Gender	F	M	F	F	F
Palsy	Left	Left	Left	Left	Left
Duration	1 year	2 years	1 month	5 months	8 months
Etiology	Vagal paraganglioma (tumour compression)	Skull base meningioma (post-excision)	Vagal paraganglioma (post-excision)	Thyroidectomy (benign)	Thyroidectomy (papillary carcinoma)
LEMG	Class III	Class V	Class V	Class III	Class IV
MUAP	+1	0	0	+1	+1
PU	+1	0	0	+3	+3
FP	0	+3	+3	0	+1
*Surgery	Nil	Hydroxyapetite IL	Collagen IL	Nil	Nil
Muscle bulk (ABSD)					
TA	Reduced 0.153	Injected	Injected	Reduced 0.121	Reduced 0.266
PCA	Reduced 0.232	Reduced 0.352	Similar 0.047	Reduced 0.167	Reduced 0.236
Signal intensity (ratio)					
TA	0.55	Injected	Injected	0.45	0.47
PCA	0.99	0.55	0.65	0.61	0.55
DCE (ratio)	PCA TA	PCA TA	PCA TA	PCA TA	PCA TA
SoE	2.85	1.37	1.51	0.97 1.92	1.66 1.71
ME	1.38	1.39	1.05	0.87 0.73	0.99 1.17
AUC	1.17	1.16	1.11	0.93 1.13	0.95 1.14
DWI Low b values	N= 4.22 P= 6.71			N= 2.28 P= 1.47	N= 2.86 P= 2.04

PU = Polyphasic unit

FP = Fibrillation potential

* : surgical intervention prior to MRI

Injected = Patients had undergone injection laryngoplasty

ABSD = Absolute value of difference

N = Normal muscle

P = Paralysed muscle

MUAP: Motor unit action potential

C. Vocal fold mobility

RESULTS: 3T-MRI cine scans demonstrated good quality of dynamic images enables measurement of vocal fold motion by measuring the glottal area (GA) (Figure 8, Figure 9). One of the patients, phonation area was not able to visualise down to distortion of the images probably due to vocal cord compensation and small bilateral laryngocoeles. In the normal volunteers, the ratio of glottal area during respiration or phonation was near to 1 with time (Figure 10, Figure 11). The mean PR across time was 0.97±0.12 and the mean RR across time was 0.94±0.1 (Table 9). This indicates that the vocal folds able to abduct and adduct equally exhibiting equal right and left GA. In UVFP patients, who all had left paralysis, the ratio of glottal area during respiration or phonation was far less or more from 1 with time (Figure 12, Figure 13). The mean PR was 0.58±0.27 and the mean RR was 1.66±0.16 (Table 9). The ratio calculated was right to left so a low PR indicates left paralysis since the left phonation area was bigger due to inability of the vocal fold to adduct and a high RR indicates that the right respiration area was bigger due to the ability of the right normal vocal fold to abduct.

Table 9: Phonation ratio (PR) and respiration ratio (RR) of volunteers and patients

Subjects	PR (SD)	RR (SD)
Volunteers	0.97±0.12	0.94±0.1
Patients	0.58±0.27	1.66±0.16

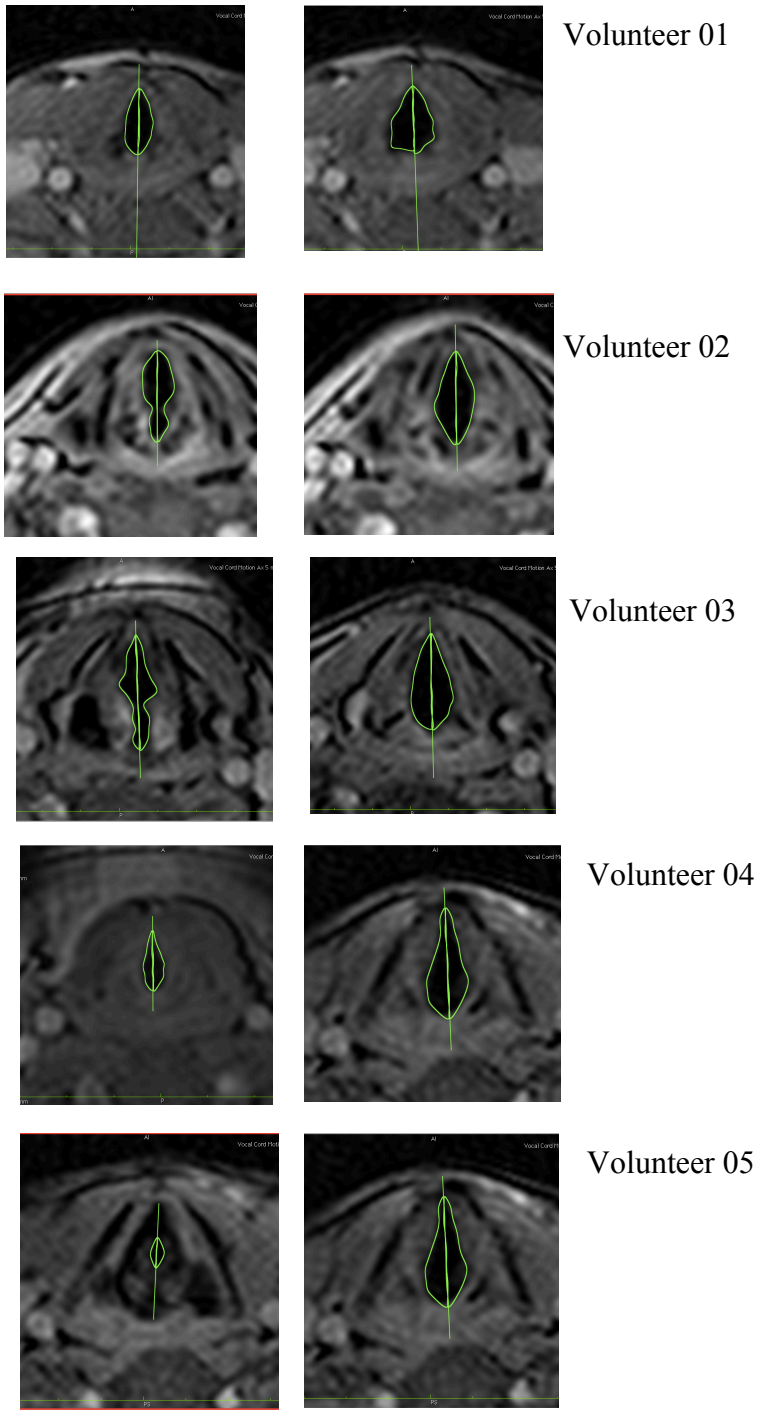


Figure 8: Images taken from cine imaging of volunteers depict phonation area (left column) and respiration area (right column). The left and right respective area was of similar size

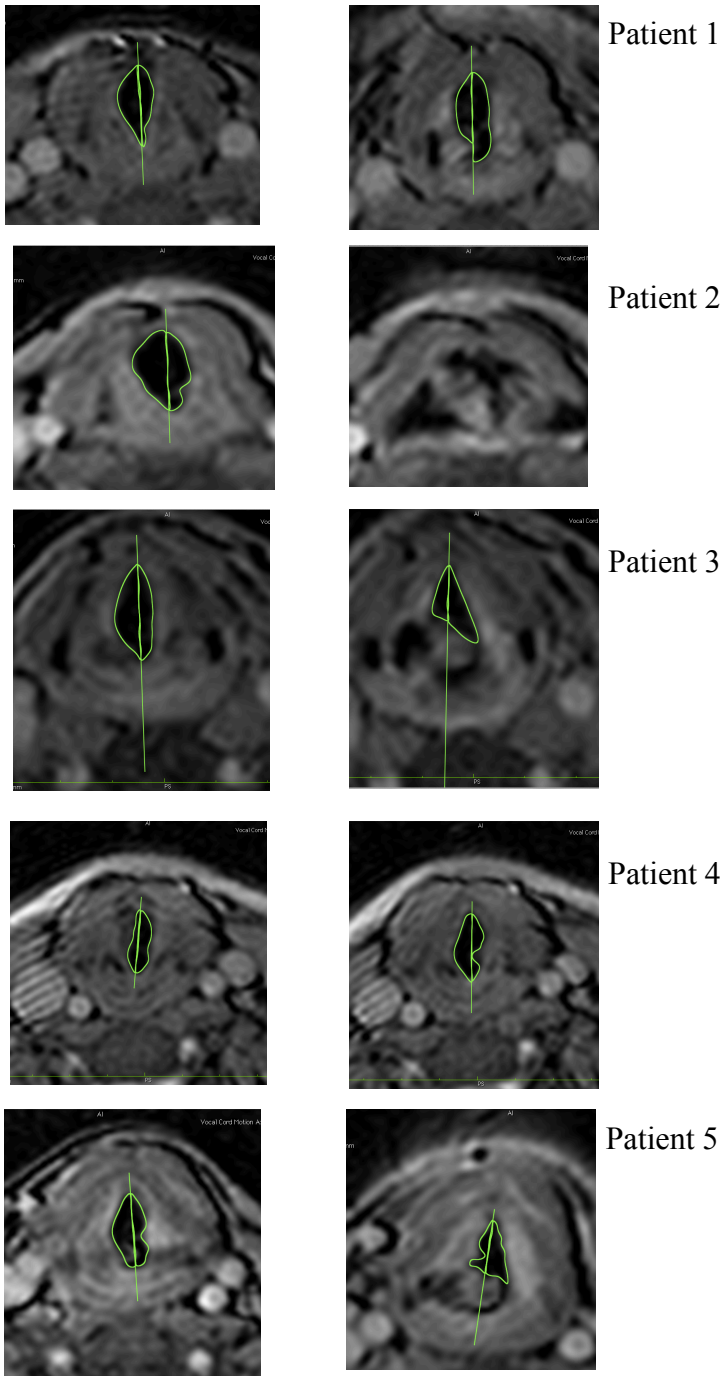
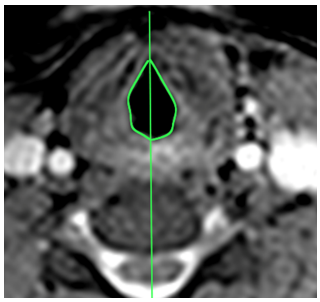


Figure 9: Images taken from cine imaging of UVFP patients depict phonation area (left column) and respiration area (right column). There was discrepancy between left and right respective area. Patient 02 phonation area was not visualised.



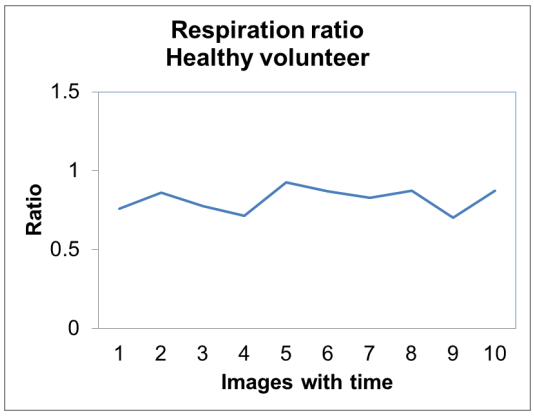
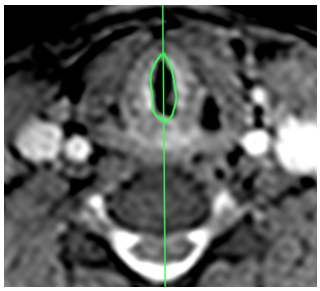


Figure 10: Respiration area right and left glottal area of a healthy volunteer. Note that the ratio across time is near to 1



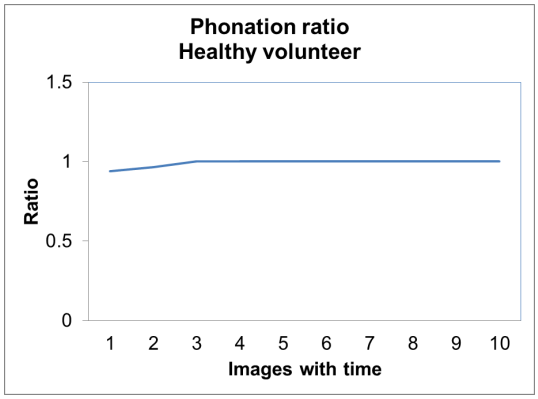
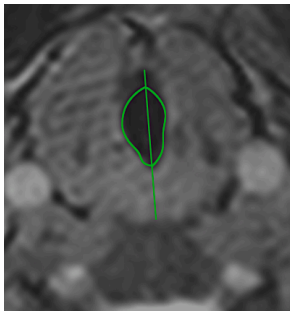


Figure 11: Phonation area right and left glottal area of a healthy volunteer. Note that the ratio across time is near to 1



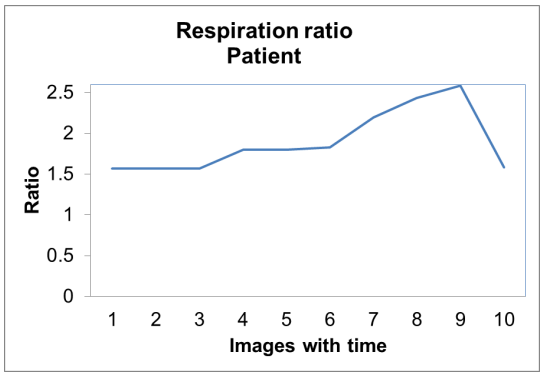
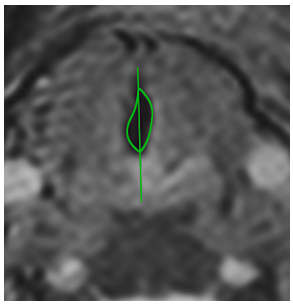


Figure 12: Respiration area of a patient with left vocal fold paralysis (arrow head). Note that the ratio across time if far more than 1



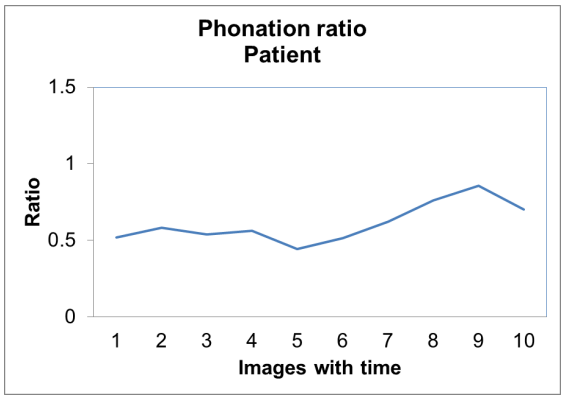


Figure 13: Phonation ratio of a patient with left vocal fold paralysis (arrow head). Note that the ratio is far less than 1 across time